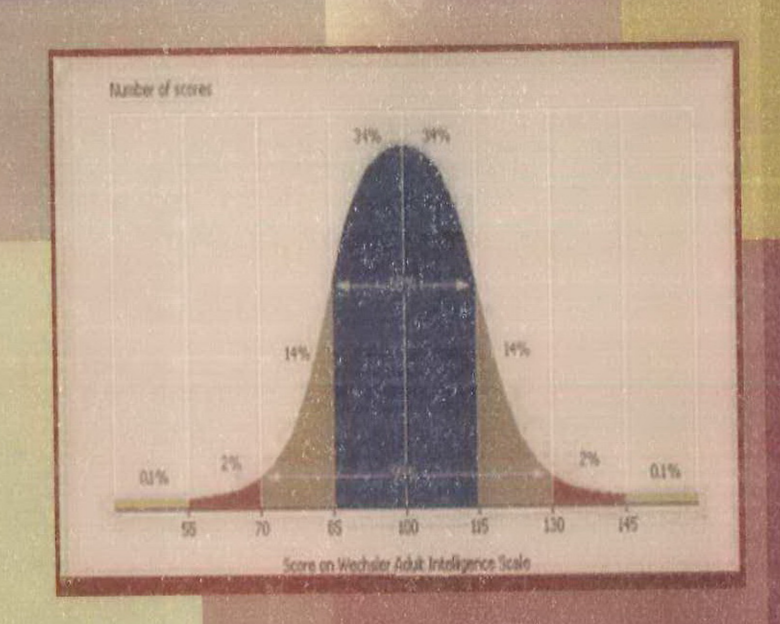
وزارة التعليم العالي والبحث العلمي
جامعة الموصل



الإحصاء الرياضي

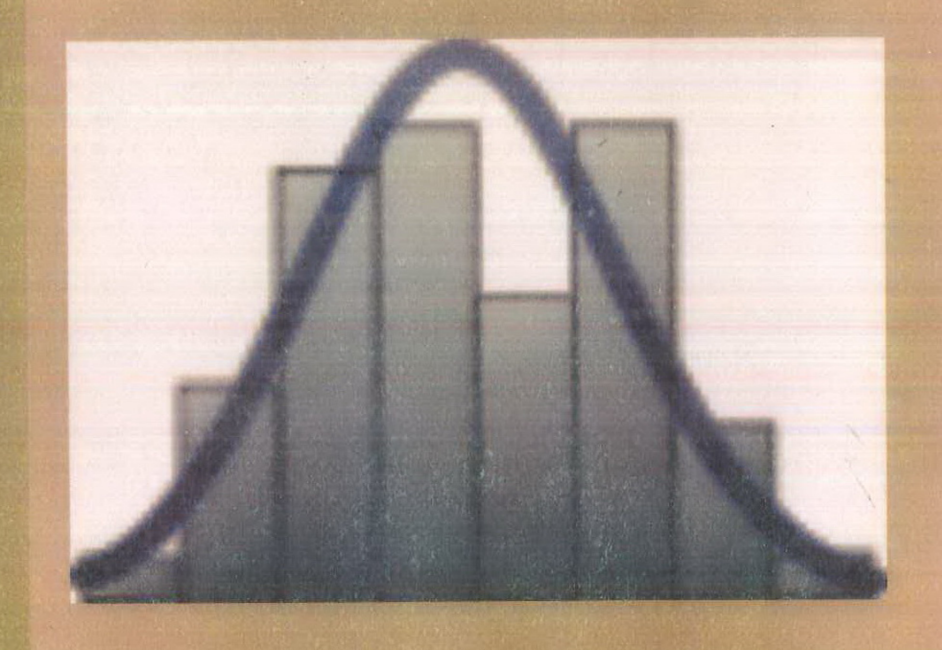


أمير حنا هرمز

الإحصاء الرياضي
أمير حنا هرمز

الإحصاء الرياضي

أمير حنا هرمز



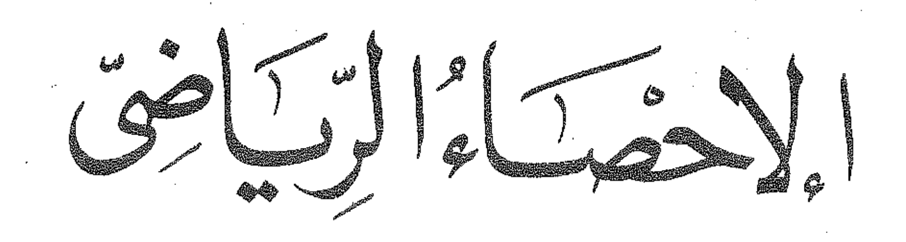


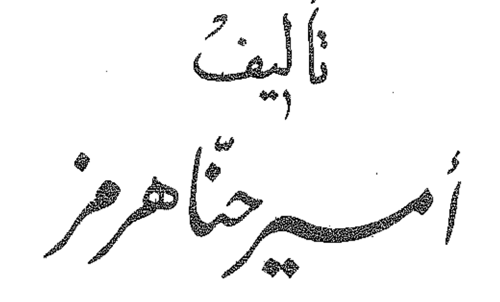


الإحصاء الرياضي

نشر وطبع وتوزيع
مديرية دار الكتب للطباعة والنشر
شارع ابن الاثير ـ الموصل
الجمهورية العراقية
هاتف ٧٦٣٢٣١
٧٦٣٢٣٥
تلكس ٨٠٩٢

وزارة التعليم العالي والبحث العلمي
جامعة الموصل





استاذ مساعد
قسم الاحصاء
كلية الادارة والاقتصاد
جامعة الموصل

*** ١٩٩٠ ***

*** المقدمة ***

يعد علم الاحصاء احدى الوسائل الهامة والحيوية في البحث العلمي ذات اصول وقواعد علمية يمكن استخدامها في ميادين العلم الاخرى التي تحتاج لاصول وقواعد وقوانين الاحصاء من خلال جمع البيانات والمعلومات اللازمة للبحث وتوظيف تلك الاصول والقواعد والقوانين في تحليل تلك البيانات بهدف الوصول الى النتائج التي يهدف لها البحث . ويعتبر علم الاحصاء بحد ذاته وسيلة وليس غاية . وذلك يعني امكانية استخدام اصول وقواعد وقوانين هذا العلم اينما وجد البحث العلمي سواء كان ذلك في مجال الاقتصاد . الزراعة . الصناعة وغيرها من المجالات . وعلم الاحصاء كبقية العلوم الاخرى شهد تطوراً سريعاً وكبيراً خلال القرنين التاسع عشر والعشرين مقترناً بتطور نظرية الاحتمالات وعلم الرياضيات . ونتيجة لهذا التطور في اصول علم الاحصاء وطرقه وقوانينه فقد ظهرت مسميات اخرى لهذا العلم اقترنت مع علوم اخرى كالاحصاء الحيوي ,الاحصاء الصناعي . الاحصاء الزراعي . الاحصاء الرياضي وغيرها .

ويعد موضوع الاحصاء الرياضي العمود الفقري للنظرية الاحصائية وأحد اركانها الهامة ذات الصلة الوثيقة بالرياضيات. ويمكن عد الاحصاء الرياضي كاحد فروع الرياضيات التطبيقية الذي يختص بتجهيز طرق واساليب وقواعد تستخدم في تحليل الظواهر ذات الطابع العددي . وسابقاً لم يكن هذا الموضوع يحمل هذا العنوان . الا انه وبمرور الزمن وزيادة عدد القواعد والنظريات الرياضية الممكنة الاستخدام في التحليل الاحصائي ادى الى عد هذه القواعد والنظريات على انها فرع من فروع الاحصاء بعنوان الاحصاء الرياضي . ان اي تقدم يحرز في مجال الرياضيات التطبيقية له وقع في رفد النظرية الاحصائية بطرق واساليب تحليلية جديدة . وهنالك مجلات علمية تخصصت في نشر البحوث والانجازات التي تعنى بالاحصاء الرياضي مثل مجلة The Annals of Mathema tical Statisics, التي تصدر عن معهد الاحصاء الرياضي في امريكا ومجلة Journal of the American Statistical Association التي تصدر عن اتحاد الاحصائيين

الاميركين ، وغيرها . مما تقدم نلاحظ ان دراسة الاحصاء الرياضي تحتاج الى المام جيد بعلم الرياضيات وخصوصاً طرق التفاضل والتكامل اضافة الى موضوع المتسلسلات النهائية واللانهائية وموضوع تقارب convergency وتباعد divergeny المتسلسلات اللانهائية كما وان لنظريات الغاية Limit theorems أهمية كبيرة في موضوع الاحصاء الرياضي وغيرها من مواضيع الرياضيات الاخرى ذات العلاقة .

لقد تم صياغة الموضوعات الواردة في هذا الكتاب بالشكل الذي يضمن سهولة فهمها واستيعابها من قبل القاريء الذي افترضنا ان يكون على المام جيد في الرياضيات وكذلك في نظرية الاحتمالات على ضوء السقررات المحددة له في دراسته الجامعية لهذين الموضوعين . هذا من ناحية ومن ناحية اخرى فقد تم الاخذ بنظر الاعتبار ان تكون موضوعات هذا الكتاب مستوفية للمقررات المحددة لموضوع الاحصاء الرياضي في اقسام الاحصاء في الجامعات العراقية مع التوسع في هذه الفقرة او تلك بهدف انماء قدرات القاريء وفق مانعتقده مناسب في سهولة فهمه لتلك الموضوعات التي تعتمد الاحصاء الرياضي كاساس لها مثل الاستدلال ، القرارات ، الدوال العشوائية وغيرها كذلك استيفاء متطلبات القاريء العام لهذا الكتاب كطلبة الدراسات العليا مثلاً ، وبناءً لحاجة القسم لوجود كتاب منهجي يفي بالمقررات الدراسية لمادة الاحصاء الرياضي لطلبة الصف الثالث احصاء . فقد صدر قرار عن مجلس كلية الادارة والاقتصاد الموقر المتخذ بالجلسة السادسة للمجلس بتاريخ ٢٥ / ١٠ / ١٩٨٨ المبلغ بالامر الاداري المرقم ٩ / ١١ / ٦٦ في ٢٣ / ١١ / ١٩٨٨ يقضي بتكليفي تأليف هذا الكتاب .

يقع هذا الكتاب في اثني عشر فصلاً . اختص الاول منها بعرض موجز لنظرية المجموعات ونظرية الاحتمالات الهدف من ذلك تذكير القاريء بهاتين النظريتين في حين تم التركيز في هذا الفصل على مفهوم المتغيرات العشوائية ودوالها . وتم تخصيص الفصل الثاني لدراسة موسعة لمفهومي التوقع الرياضي والدوال المولدة للعزوم لاهميتها في فهم الكثير من الفقرات اللاحقة لهذا الفصل . وبغية استكمال خصائص دوال المتغيرات العشوائية فقد تم تخصيص الفصل الثالث لدراسة اهم المقاييس الاخرى ذات العلاقة بالتوزيع الاحتمالي كالمنوال والوسيط وغيرها . وتركزت دراستنا في الفصل الرابع في عرض واف لمفهوم التوزيعات المشتركة

والشرطية وخصائص هذه التوزيعات واهم الامور ذات العلاقة بها . اما الفصل الخامس فقد اختص بدراسة شاملة وافية لاهم التوزيعات الاحتمالية المتقطعة الشائعة الاستخدام في تطبيقات النظرية الاحصائية . في حين اختص الفصل السادس بدراسة شاملة لاهم التوزيعات المستمرة الشائعة . واختص الفصل السابع في دراسة توزيعات دوال المتغيرات العشوائية مع استعراض لاهم الطرق المتبعة في استنتاج توزيعات هذه الدوال . اما الفصل الثامن فقد خصص لدراسة موضوعي المعاينة والتوزيعات المقيدة مع عرض واف لقانون الاعداد الكبيرة ومبرهنة الغاية المركزية . ونظراً لاهمية توزيعات المعاينة في تطبيقات النظرية الاحصائية وخصوصاً في موضوعي اختبار الفرضيات وفترات الثقة فقد تم تخصيص الفصل التاسع لدراسة اهم هذه التوزيعات وبشكل مفصل مع عرض لاهم استخداماتها . وتم تخصيص الفصل العاشر لدراسة مفهوم الاحصاءات المرتبة وتوزيعات دوال هذه الاحصاءات . وبغية تعريف القارىء بنظرية التقدير واختبار الفرضيات فقد تم تخصيص الفصل الحادي عشر لدراسة موجزة لنظرية التقدير بنقطة والتقدير بفترة . واخيراً فقد تم تخصيص الفصل الثاني عشر لدراسة موجزة لاختبار الفرضيات .

لقد تم الاخذ بنظر الاعتبار وبهدف توسيع مدارك القارىء تعزيز كل فصل من فصول هذا الكتاب بمجموعة من الامثلة التوضيحية اضافة الى مجاميع من التمارين موزعة على فقرات كل فصل او في نهاية كل فصل

وختاما يقتضي واجب الوفاء ان اتقدم بوافر الشكر والامتنان الى الدكتور عادل فليح العلي عميد الكلية لتشجيعه تأليفي هذا الكتاب متمنياً له دوام الموفقية . كذلك اتقدم باسمى آيات الشكر والتقدير لكل من الدكتور عبد الجبار البرهاوي والدكتور سعد اسحق عطية والدكتورة برلنتي جميل شمعون لما بذلوه من جهود قيمة في مراجعة مسودات الكتاب وتسجيل ملاحظاتهم بشأنها واسجل شكري وتقديري لسادة رئيس واعضاء مجلس قسم الاحصاء لما قدموه من دعم معنوي طيلة فترة تأليفي هذا الكتاب . وتقتضي الامانة العلمية ان اسجل وافر شكري وتقديري لكل من الاستاذ الدكتور قبيس سعيد الفهادي المقوم العلمي للكتاب والدكتور عبد الوهاب العدواني / رئيس قسم اللغة العربية / كلية الاداب المقوم اللغوي للكتاب لمراجعتهم مسودات الكتاب وتسجيل ملاحظاتهم القيمة بشأنها متمنيا لهما دوام الموفقية .

ولكافة العاملين في مديرية مطبعة التعليم العالي في الموصل اسجل اسمى آيات الشكر والتقدير للجهود القيمة التي بذلوها في اخراج الكتاب متمنيا لهم الموفقية في عملهم .

ارجو ان اكون قد وفقت في اخراج هذا الكتاب بالشكل الذي يفي باحتياجات القاريء العزيز وتحقيق الاهداف المتوخاة منه خدمة لعراقنا العزيز . وأمل من زملائي الافاضل مدرسي مادة الاحصاء الرياضي موافاتي بملاحظاتهم القيمة لاغناء الكتاب في طبعته القادمة ومن الله التوفيق

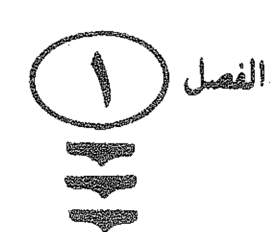
المؤلف

أمير حنا هرمز العوصه جي

١٩٩٠

* * * المحتويات * * *





مقدمة في نظرية الاحتمالات

الفصل الاول

مقدمة في نظرية الاحتمالات

ان الاحصاء الرياضي في الحقيقة هو استكمال لنظرية الاحتمالات probability theory وهذا يعني ان الامر يتطلب الماماً جيداً بنظرية الاحتمالات التي هي الاخرى تستخدم نظرية المجموعات Set theory والعمليات الخاصة بها . وبهدف تذكير القاريء بهاتين النظريتين فاننا سوف نوجزهما بالفقرتين التاليتين :

١ ــ ١ : نظرية المجموعات Set theory

تعرف المجموعة بأنها جمع لاشياء معينة وان هذه الاشياء قد تكون اعداد او كميات (اوا ي شيء آخر) تشترك بصفة معينة . على سبيل المثال مجموعة الاعداد الزوجية المحصورة بين العددين 1,11 . مجموعة الحروف الانكليزية الصغيرة ، اي a, b, c, ..., z . قائمة باسماء طلبة صف معين . مجموعة المحافظات العراقية . وغيرها من الامثلة . وعادة يرمز للمجموعة باحد الاحرف الانكليزية الكبيرة مثل A, B, ... ويطلق على كل شيء معرف في المجموعة بالعنصر element وهذا يعني ان A. عبارة عن مجموعة من العناصر . غالباً مايتم حصرها بقوسين من الشكل { } للدلالة على المجموعة فمثلاً مجموعة الاعداد الزوجية المحصورة بين 1,11 هــي $A = \{2, 4, 6, 8, 10\}$ ومجمـوعة الـحروف الانكلـيزية الصغيرة هي $B = \{a, b, c, ..., z\}$. ومجموعة اسماء طلبة صف معين هي { حيث x طالب في الصف المعني : x } = C . ان ترتيب العناصر داخل المجموعة من حيث سبق عنصر على آخر لايعني شيئاً ولا يؤثر على سياق تعريف المجموعة . ويرمز لعائدية اي عنصر لمجموعة معينة بالرمز belonge to ∈ . فاذا كان x عنصر معرف في A عندئذٍ $x \in A$ وفيما يلي بعض التعاريف والمصطلحات التي سترد في هذه الفقرة :

١ ـ ١ ـ ١ : تعاريف ومصطلحات :

ان مجموعة كل العناصر التابعة للمشكلة قيد الدراسة تسمى بالمجموعة الشاملة (او الفضاء) . universal set (or space) . وغالباً مايرمز لهذه المجموعة بالرمز S او Ω . وعلى فرض ان $S = \{a, b, c, ..., z\}$ عندئذٍ فان :

١ ـ المجموعة الجزئية : Sub set

افرض ان $A = \{a, b, c, d\}$ عندئذٍ يقال ان A مجموعة جزئية من S بسبب احتواء A على بعض العناصر المعرفة في S . ويعبر عن ذلك بالقول ان A محتواة في S ، ويرمز لذلك بالشكل $A \subset S$ واذا كانت $B = \{a,c\}$ عندئذٍ فان B مجموعة جزئية من A (وكذلك من S) بسبب احتوائها على بعض العناصر المعرفة في A . وهذا يعني ان $B \subset A$

٢ ـ المجموعات المتكافئة Equivalent sets

يقال المجموعتين C,D متكافئتان اذا كانت كل منهما محتواه في الاخرى اي ان $C \subseteq D$ وان $C \subset D$ وعندئذٍ يقال ان $C = D$ فاذا كانت $C = \{e, f, g\}$, $D = \{e, f, g\}$ نلاحظ ان C محتواة في D وكذلك D محتواة في C

٣ ـ المجموعة الخالية Empty set (null set).

يقال ان مجموعة معينة هي مجموعة خالية اذا كانت هذه المجموعة لا تحتوي على اي عنصر . ويرمز لهذه المجموعة بالرمز ϕ

٤ ـ متممة المجموعة Complement of a set

لتكن $F = \{a, b, c, d, e, f\}$ مجموعة جزئية من S . عندئذٍ فان المجموعة المتممة الى F . ولتكن F^c . تمثل تلك المجموعة التي تحتوي على كافة العناصر المعرفة في S الغير معرفة في F . وهذا يعني ان $F^c = S - F$ اي ان $F^c = \{g, h, i, ..., z\}$

٥ ـ فضلة المجموعة : Set difference:

لتكن { G ={ e, f, g, h, i, j } , H = { g, h, j, k, l, m مجموعتين معرفتين فيS. عندئذٍ تعرف فضلة H على G بانها مجموعة العناصر المعرفة في G الغير معرفة في H . ويرمز للفضلة بالرمز G/H . وهذا يعني ان { G / H = { e, f, i . كذلك فان H / G={ K, L, m }

٦ ـ اتحاد المجموعات Union of the sets

لتكن{ I={ g, h, p, q, r, s }, J={ a, b, h, q, r مجموعتين معرفتين في S . عندئذٍ تعرف المجموعة الناتجة عن اتحاد I مع J بانها المجموعة المؤلفة من كافة العناصر المعرفة في I او في J او في كليهما . ويرمز لاتحاد مجموعتين بالرمز $\bigcup$. وهذا يعني ان { I $\bigcup$ J = { a, b, g, h, p, q, r, s

٧ ـ تقاطع المجموعات Intersection of the sets

لتكن{ K = { a, b, c, d, p } , L = { d, p, q, r, s مجموعتين معرفتين فيS .عندئذٍ تعرف المجموعة الناتجة عن تقاطع K مع L بأنها المجموعة المؤلفة من العناصر المعرفة في K وفي ذات الوقت معرفة في L . ويرمز لتقاطع مجموعتين بالرمز $\bigcap$. وهذا يعني { K $\bigcap$ L = { d,p وفي حالة عدم وجود عنصر واحد على الاقل معرف في كلا المجموعتين عندئذٍ فان التقاطع هو مجموعة خالية ϕ

ان العمليات التي تجرى على المجموعات كالاتحاد والتقاطع محكمة بقوانين وبديهيات تفسر العلاقات بين المجموعات . وعلى فرض ان A, B, C ثلاث مجموعات . عندئذٍ فان هذه القوانين والبديهيات تنص على ما يلي :

أ ـ قانون الا بدال Commutative law

ان $A \bigcap B = B \bigcap A \quad , \quad A \bigcup B = B \bigcup A$

ب ـ قانون الترتيب Associative law

ان $A \bigcap (B \bigcap C) = (A \bigcap B) \bigcap C \quad , A \bigcup (B \bigcup C) = (A \bigcup B) \bigcup C$

جـ ـ قانون التوزيع Distributive law

ان $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

وان $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

د ـ ان $(A^c)^c = A$ أي أن متممة مجموعة مثل A تمثل المجموعة A

هـ ـ قانونا اللانمو Idempotent law

اذا كانت A مجموعة جزئية من S عندئذٍ فان

$$A \cup S = S , A \cap S = A$$

وان
$$A \cup \phi = A , A \cap \phi = \phi$$

و ـ لتكن A مجموعة جزئية من S . عندئذٍ فان

$$A \cup A^c = S , A \cap A^c = \phi$$

وان
$$A \cup A = A \cap A = A$$

ز ـ قانونا دي موركان De Morgans' laws

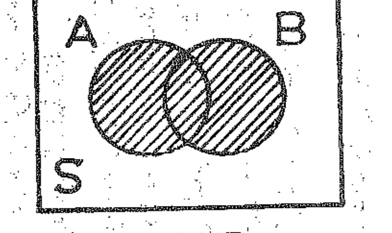
لتكن A,B مجوعتين معرفتين في S . عندئذٍ

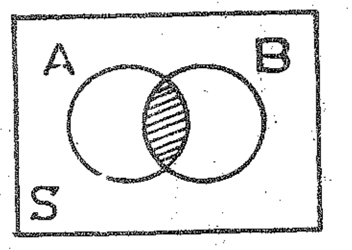
$$(A \cup B)^c = A^c \cap B^c , (A \cap B)^c = A^c \cup B^c$$

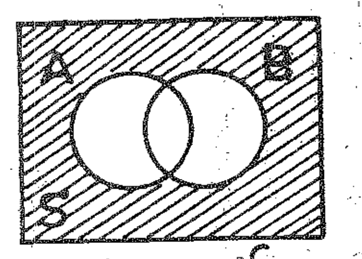
ح ـ ان فضلة A على B تمثل المجموعة الناتجة من تقاطع A مع متممة B

اي ان $B / A = A \cap B^c$

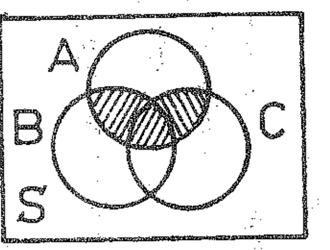
وغالبا مايتم استخدام مخططات توضيحية تسمى مخططات ڤين venn diagrams الهدف منها اعطاء صورة واضحة عن العمليات التي تجرى على المجموعات . وفيما يلي بعض من هذه المخططات التي توضح بعض مما سبق:







$(A \cap B)^C$



$A \cup (B \cap C)$

الشكل (١ ـ ١) ، مخططات ڤين .

١ ـ ١ ـ ٢ : الشكل المختصر في التعبير عن المجموعة

تصادفنا في احوال كثيرة مجموعات لايمكن حصر كافة عناصرها مما يستوجب الامر البحث عن شكل آخر يمكن من خلاله التعبير عن هذا النوع من المجموعات. فمثلاً لو كانت A تمثل مجموعة كافة الاعداد الحقيقية غير السالبة فأنه من الصعوبة حصر كافة عناصر A. علية وبفرض ان x يمثل اي عنصر في A (اي ان $x \in A$) عندئذٍ يمكن التعبير عن المجموعة A بالشكل $A = \{ x : x \geq 0 \}$ واذا كانت B مجموعة كافة الاعداد الصحيحة الفردية الموجبة عندئذٍ يمكن التعبير عن B بالشكل $B = \{ x :$ عدد صحيح فردي موجب $x \}$ ويلاحظ من هذين المثالين ان الرمز x هو رمز مميز لاي عنصر ينتمي للمجموعة دون الحاجة الى ذكر تفاصيل المجموعة.

وفيما يلي بعض التعابير والمصطلحات عن المجموعات والتي يتم الاستفادة منها عند دراستنا لموضوع المتغيرات العشوائية ودوالها في الفقرة (١ ـ ٣).

١ ـ المجموعة المنتهية : Finite set

يقال ان المجموعة A مجموعة منتهية اذا كان عدد عناصرها مساوٍ الى n ، حيث n عدد محدود ، او انها مجموعة خالية . وفي غير ذلك يقال ان A مجموعة غير منتهية infinite set . فمثلاً المجموعة $A = \{ 2, 4, 6, 8 \}$ هي مجموعة منتهية بسبب احتوائها على عدد محدود من العناصر . في حين ان المجموعة $B = \{ 1, 2, 3, \ldots \}$ والمجموعة $C = \{ x : 0 < x < 4 \}$ مجموعتين غير منتهيتين بسبب احتوائها على عدد غير منته (غير محدود) من العناصر.

٢ . المجموعة القابلة للعد : Countable set

يقال ان المجموعة A قابله للعد اذا امكن عد (ملاحظة) عناصر هذه المجموعة . وفي غير ذلك يقال ان A مجموعة غير قابلة للعد uncountable set . فمثلا المجموعة $A = \{ 1, 2, 3, \ldots \}$ والمجموعة $B = \{ 2, 4, 6, 8 \}$ مجموعتان قابلتان للعد . في حين ان المجموعة $C = \{ x : 0 < x < 6 \}$ والمجموعة $D = \{ x : 0 < x < \infty \}$ مجموعتان غير قابلتين للعد .

مثال (١): افرض $\Omega = \{x : x = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ وان كل من A, B, C, D مجموعة جزئية معرفة في Ω بحيث ان $A = \{x : x < 5\}$ ، $B = \{x : x \geq 4\}$ ، $C = \{x :$ عدد زوجي $x\}$ ، $D = \{x :$ عدد فردي $x\}$ عندئذٍ فان :

$$A \cup B = \{x : x = 2, 3, \ldots, 12\} = \Omega$$
$$C \cup D = \{x : x = 2, 3, \ldots, 12\} = \Omega$$
$$B \cup D = \{x : x = 3, 4, \ldots, 12\}$$
$$A \cap B = \{x : x = 4\}, A \cap C = \{x : x = 2, 4\}, A \cap D = \{x : x = 3\}$$
$$B \cap C = \{x : x = 4, 6, 8, 10, 12\}, B \cap D = \{x : x = 5, 7, 9, 11\}$$
$$A^c \cap B = \{x : x \geq 5\} \cap \{x : x \geq 4\} = \{x : x = 5, 6, 7, 8, 9, 10, 11, 12\}$$
$$A^c \cap C = \{x : x \geq 5\} \cap \{x : \text{عدد زوجي } x\} = \{x : x = 6, 8, 10, 12\}$$
$$(A \cup D)^c = \{x : x = 6, 8, 10, 12\}, (B \cup C)^c = \{x : x = 2, 3\}$$
$$(B \cap D)^c = \{x : x = 2, 3, 4, 6, 8, 10, 12\}$$
$$(A \cup B \cup C)^c = \phi, (A \cap B \cap C)^c = \{x : x = 2, 3, 5, 6, \ldots, 12\}$$
$$((A \cup B) \cap D)^c = (\Omega \cap D)^c = C$$

مثال (٢) : افرض ان $\Omega = \{x : x \geq 0\}$ وان كل من A, B, C مجموعة جزئية في Ω بحيث ان $A = \{x : x > 5\}, B = \{x, 2 < x < 7\}, C = \{x : x < 6\}$
عندئذٍ فان .

$$A^c = \{x : x > 5\}^c = \{x : x \leq 5\}$$
$$B^c = \{x : 2 < x < 7\}^c = \{x : x \leq 2, x \geq 7\}$$
$$(A \cup B)^c = \{x : x \leq 2\}, (B \cup C)^c = \{x : x \geq 7\}$$
$$(A \cap B)^c = \{x : x \leq 5, x \geq 7\}$$
$$(A \cup B) \cap C = \{x : 2 < x \leq 6\}, (A \cup C) \cap B = B$$
$$(B \cap C) \cup A = \{x : x > 2\}, A \cap B \cap C = \{x : 5 < x \leq 6\}$$

مثال (٣) : افرض $\Omega = \{x : -\infty < x < \infty\}$ وان $A = \{x : x < 8\}$ و $B = \{x :$ عدد صحيح موجب $x\}$ عندئذٍ فان :

$$A \cap B = \{x : x = 1, 2, 3, \ldots, 8\}$$

لاحظ هنا انه بالرغم من ان A مجموعة غير منتهية وغير قابلة للعد وان B غير منتهية قابله للعد الا ان تقاطع A مع B يمثل مجموعة منتهية قابلة للعد . كذلك فان .

$$A \cup B = \{ x : x \leq 8 , x = 9 , 10 , \ldots \}$$

$A^c \cap B = \{ x : x \geq 9$ حيث x عدد صحيح $\}$

١ ـ ٢ : نظرية الاحتمالات probability theory

بعد استعراضنا وبشكل موجز لنظرية المجموعات ناتي في هذه الفقرة الى شرح موجز ايضاً لنظرية الاحتمالات ومن خلال الفقرات التالية .

١ ـ ٢ ـ ١ : فضاء العينة Sample space

ان مجموعة النتائج الممكنة لتجربة عشوائية معينة تسمى « فضاء العينة » . حيث سبق وان اشرنا لهذا المفهوم في الفقرة (١ ـ ١ ـ ١) ورمزنا له بالرمز Ω . ويقصد بالتجربة العشوائية بانها تلك التجربة التي لايمكن التعرف على نتائجها الا بعد تنفيذها . فمثلاً في تجربة رمي قطعة نقود منتظمة نلاحظ وجود نتيجتين ممكنتين فقط هما ظهور صورة أو ظهور كتابة بعد رمي القطعة واستقرارها . وفي تجربة رمي زهر نرد منتظم نلاحظ وجود ست نتائج ممكنة فقط هي 6 , 5 , 4 , 3 , 2 , 1 بعد رمي زهر النرد واستقراره . اما في تجربة رمي زهري نرد منتظمتين فان عدد النتائج الممكنة هي $36 = 6^2$ نتيجة . ان كل « نتيجة outcome » ممكنة لتجربة عشوائية هي في الحقيقة عنصر ينتمي لمجموعة النتائج الممكنة Ω .

١ ـ ٢ ـ ٢ : الحوادث Events

ان اية مجموعة جزئية معرفة في Ω تسمى « الحادثة » . ففي تجربة رمي زهر نرد فان المجموعة الجزئية $A = \{ x : x = 2 , 4 , 6 \}$ تمثل حادثة معرفة في Ω . وفي تجربة رمي زهري نرد وبفرض ان x يمثل عدد النقاط الظاهرة على وجه الزهر الاول و y تمثل عدد النقاط الظاهرة على وجه الزهر الثاني فان المجموعة الجزئية $B = \{ (x,y) : 6 \leq x + y \leq 9 \}$ تمثل حادثة معرفة في Ω وان

$C = \{(x,y) : x + y \geq 7\}$ تمثل حادثة اخرى معرفة في Ω ، كذلك فان $B \cup C = \{(x,y) : 6 \leq x + y \leq 12\}$ هي ايضاً حادثة معرفة في Ω وكذلك الحادثة $B \cap C = \{(x,y) : 7 \leq x + y \leq 9\}$.

ان فئة كل الحوادث الممكنة التعريف في تجربة عشوائية تسمى « فضاء الحادثة "eventspace" . كذلك اذا كانت A , B حادثتين معرفتين في Ω عندئذٍ يقال ان هاتين الحادثتين متنافيتان mutually exclusive events اذا كان وقوع احداهما يمنع وقوع الاخرى وهذا يعني ان $A \cap B = \phi$ فمثلاً حادثة ظهور صوره عند رمي قطعة نقود منتظمة تمثل حادثة متنافية مع حادثة ظهور كتابة اي ان كل منهما يمنع تماماً ظهور الاخرى . كذلك يقال ان نتائج تجربة عشوائية معينة تمتلك نفس الفرصة في الوقوع equally likely outcomes اذا لم يكن هنالك سبب لتفضيل نتيجة على أخرى . فمثلاً عند رمي قطعة نقود منتظمة مرة واحدة فان ظهور الصورة او ظهور الكتابة هي ذات نفس الفرصة في الوقوع ، وعند رمي زهر نرد مرة واحدة نلاحظ ان ظهور اي وجه من أوجه الزهر يمتلك فرصة مماثلة لظهور اي وجه آخر .

١ ــ ٢ ــ ٣ : تعريف الاحتمال Definition of probability

افرض ان عدد النتائج الممكنة في تجربة عشوائية هو n من النتائج المتنافية وذات نفس الفرصة في الوقوع وان $m , m \leq n$ من هذه النتائج ممكنة الوقوع في حادثة معينة مثل A معرفة في Ω عندئذٍ فان « احتمال حدوث A » . ويرمز له بالشكل $P_r(A)$ أو $P(A)$. معرف بالآتي :

$$P_r(A) = P(A) = P = \frac{m}{n}$$

وغالباً مايشار الى p على انه احتمال نجاح وقوع الحادثة A في حين ان احتمال فشل وقوع A هو $q = \frac{n-m}{n}$ وهذا يعني ان $q = 1 - p$. فمثلاً احتمال ظهور الوجه الذي يحمل نقاط عددها 4 في تجربه رمي زهر نرد هو

$P_r(A) = P_r(x=4) = \frac{1}{6}$. كذلك فان احتمال ظهور الوجه 4 أو الوجه 6 هو

$$P_r(B) = P_r(x = 4 \text{ or } 6) = \frac{2}{6} = \frac{1}{3}.$$

يتضح من التعريف اعلاه انه يشترط في حساب احتمال حدوث A، ان يكون عدد النتائج الممكنة n عدداً محدوداً finite وفي غير ذلك لايمكن حساب الاحتمال . ان هذا التعريف للاحتمال هو تعريف كلاسيكي . وهنالك تعريف آخر له يسمى « الاحتمال الاحصائي » او « الاحتمال التجريبي » الذي يستند الى تفسير التكرار النسبي relative frequency لتجربة معينة كررت بعدد من المرات المتجانسة والمتشابهة وان عدد النتائج الممكنة للتجربة غير محدود . وبفرض ان عدد مرات تكرار التجربة كبير عندئذٍ يمكن تعريف الاحتمال على انه غاية النسبة ما بين عدد مرات حدوث الحادثة A وعدد مرات التكرار . وبفرض ان n تمثل عدد مرات تكرار التجربة وان m تمثل عدد مرات وقوع الحادثة A عندئذٍ فان احتمال وقوع A هو

$$P_r(A) = \lim_{n \to \infty} \frac{m}{n}$$

فمثلاً لو كانت A تمثل الحصول على الوجه 4 في تجربة رمي زهر نرد وبفرض ان هذه التجربة كررت بعدد كبير من المرات عندئذٍ ووفق هذا التعريف فان النسبة ما بين عدد مرات ظهور الوجه 4 وعدد مرات تكرار التجربة يقترب من $\frac{1}{6}$. اي انه عندما $n \to \infty$ فان $P_r(A) = P_r(x = 4) \to \frac{1}{6}$.

١ ـ ٢ ـ ٤ : بديهيات الاحتمال Axioms of probability

ان مقياس الاحتمال P لحادثة معينة مثل A معرفة في Ω يمثل دالة ذات قيمة حقيقية real – valued function معرفة في الفترة $[0, 1]$ وتحقق الشروط التالية :

١ ـ ان $0 \leq P \leq 1$. وهذا يعني ان احتمال وقوع اي حادثة مثل A معرفة في Ω تتراوح قيمته ما بين الصفر والواحد . وعندما تكون $P = 1$ فذلك يعني ان الحادثة A مؤكدة الوقوع . كالقول « ماهو احتمال الحصول على كرة حمراء من صندوق يحتوي على خمس كرات حمراء متجانسة ؟ » ان هذا الاحتمال مساوٍ للواحد بسبب ان كافة الكرات حمراء اللون . وعندما تكون $P = 0$ فذلك يعني استحالة وقوع A . كالقول « ماهو احتمال الحصول على كرة سوداء من صندوق يحتوي

على خمس كرات حمراء ؟ » ان هذا الاحتمال مساوٍ للصفر بسبب عدم وجود كرة سوداء في الصندوق .

٢ ـ ان $P_r(\Omega)=1$. وهذا يعني ان احتمال حدوث الفضاء مساوٍ للواحد كالقول « ماهو احتمال الحصول على وجه يحمل على الاقل نقطة واحدة في تجربة رمي زهر نرد ؟ » واضح ان هذا الاحتمال مساوٍ للواحد بسبب ان الحادثة هنا تمثل ظهور أي وجه من اوجه الزهر .

٣ ـ ان احتمال اتحاد عدد من الحوادث المتنافية المعرفة في Ω مثل $A_1, A_2, \ldots$ هو مجموع احتمالات هذه الحوادث . اي ان

$$P_r(A_1 \cup A_2 \cup \ldots) = \sum_i P_r(A_i)$$

واذا كان عدد هذه الحوادث محدود ومساوٍ الى n مثلاً فان

$$P_r(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} P_r(A_i)$$

هذه البديهيات الثلاث تسمى بديهيات الاحتمال التي من خلالها يمكن تحديد مفهوم الاحتمال . واستناداً لهذه البديهيات ولاي حادثتين مثل A, B معرفتين في Ω فانه :

ـ اذا كانت $A \subset B$ عندئذٍ $P_r(A) \leq P_r(B)$.

ـ اذا كانت $A \cap B = \phi$ عندئذٍ $P_r(A \cap B) = P_r(\phi) = 0$

ـ اذا كانت $B = A^c$ عندئذٍ $P_r(A \cup A^c) = P_r(A) + P_r(A^c) = 1$

١ ـ ٢ ـ ٥ : قاعدة جمع الاحتمالات Addition rule of probabilites

لتكن A, B أي حادثتين معرفتين في Ω فان

$$P_r(A \cup B) = P_r(A) + P_r(B) - P_r(A \cap B)$$

واذا كانت A, B متنافيتين فان $P_r(A \cup B) = P_r(A) + P_r(B)$ بسبب ان $P_r(A \cap B) = 0$

مثال (١) : في تجربة رمي زهر نرد فان :

$$P_r(x: x \text{ عدد زوجي او } x \geq 4) = P_r(x = 2, 4, 6) + P_r(x = 4, 5, 6)$$

$$- P_r(x = 4, 6) = \frac{2}{3}$$

كذلك فان

$$P_r(x \leq 2 \text{ او } x \geq 3) = P_r(x \leq 2) + P_r(x \geq 4)$$

$$= \frac{2}{6} + \frac{3}{6} = \frac{5}{6}$$

١ ـ ٢ ـ ٦ : قاعدة ضرب الاحتمالات

Multiplication rule of probabilites

افرض ان A , B اي حادثتين معرفتين في Ω . عندئذٍ يقال ان A مستقلة عن B اذا وفقط اذا كان $P_r(A \cap B) = P_r(A).P_r(B)$ وفي غير ذلك يقال ان A , B معتمدتان وهذا يعني ان وقوع A مرهون بوقوع B وان وقوع B مرهون بوقوع A ومن ذلك نستشف ان وقوع اي من هاتين الحادثتين مشروط بوقوع الاخرى . فاذا فرضنا ان A سوف تقع علماً ان B قد وقعت . ويرمز لذلك بالشكل $A | B$. فان احتمال وقوع A في هذه الحالة هو $P_r(A | B)$ وهذا الاحتمال يسمى « الاحتمال الشرْطي » conditional probability اي احتمال وقوع A علماً ان B قد وقعت . وعندئذ فان احتمال وقوع A , B معاً (اي احتمال التقاطع) هو :

$$P_r(A \cap B) = P_r(A | B) . P_r(B)$$

وعلى هذا الاساس يعرف الاحتمال الشرْطي للحادثة A علماً ان B قد وقعت فعلاً كالآتي :

$$P_r(B|A) = \frac{P_r(A \cap B)}{P_r(A)}, P_r(A) > 0, P_r(A|B) = \frac{P_r(A \cap B)}{P_r(B)}, P_r(B) > 0$$

ويلاحظ مما تقدم اذا كانت B مستقلة عن A فان

$$P_r(A | B) = P_r(A) \; , \; P_r(B | A) = P_r(B)$$

مثال (٢) : في تجربة رمي زهري نرد احدهما احمر اللون والثاني ابيض وان x تمثل عدد النقاط الظاهرة على وجه الزهر الاحمر وان y عدد النقاط الظاهرة على وجه الزهر الابيض بعد استقرارهما. ولتكن $A = \{ (x, y) : x + y \leq 4 \}$ وان $B = \{ (x, y) : x = 1 \text{ or } y = 1 \}$ عندئذٍ فان :

$$A \cap B = \{ (x, y) : (x, y) = (1, 1), (1, 2), (1, 3), (2, 1), (3, 1) \}$$

وهذا يعني ان $P_r(A \cap B) = \frac{5}{36}$. كذلك فان :

$$B = \{ (x, y) : (x, y) = (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1), (3, 1), (4, 1), (5, 1), (6, 1) \}$$

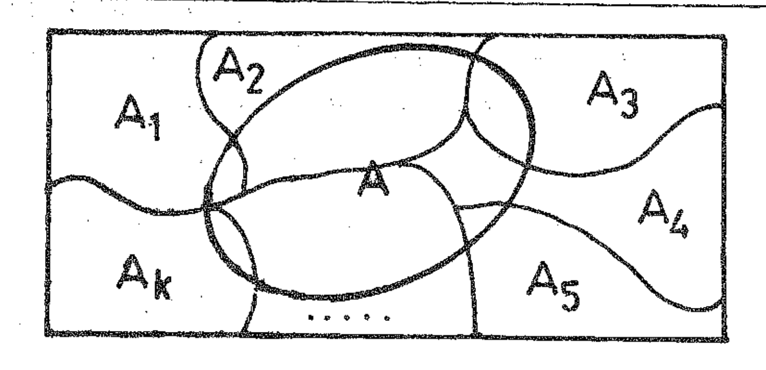
فاذن $P_r(B) = \frac{11}{36}$. وهذا يعني ان :

$$P_r(A \mid B) = \frac{P_r(A \cap B)}{P_r(B)} = \frac{5}{11}$$

وحيث اننا بصدد التكلم عن الاحتمال الشرطي نرى من الضروري ذكر بعض الشيء عن نظرية بيز Bays theorem التي تنص على ما يلي :

افرض ان Ω يمثل فضاء العينة لتجربة معينة وانه امكن تجزئة هذا الفضاء الى عدد من المجموعات الجزئية (حوادث) المنفصلة مثل $A_1, A_2, \ldots A_k$ بحيث ان $\bigcup A_i = \Omega$ وان $P_r(A_i) > 0$. عندئذٍ ولاية مجموعة معينة مثل A معرفة في Ω بحيث ان $0 < P_r(A)$ فان :

$$P_r(A_i \mid A) = \frac{P_r(A_i) \cdot P_r(A \mid A_i)}{\sum_{i=1}^{k} P_r(A_i) \cdot P_r(A \mid A_i)}$$



الشكل (١ ــ ٢) : تجزئة الفضاء Ω

مثال (٣) : يوجد ثلاثة صناديق هي a , b , c الصندوق a يحتوي على ست كرات ثلاث منها حمراء واثنين سوداء والاخرى بيضاء . الصندوق b يحتوي على اربع كرات واحدة حمراء واخرى سوداء والبقية بيضاء الصندرق c يحتوي على اثنتي عشر كرة ثلاث منها حمراء وخمس سوداء والبقية بيضاء . اختبر احد هذه الصناديق عشوائيا وسحبت منه كرتان ولوحظ ان احدهما بيضاء والاخرى حمراء . ماهو احتمال ان هاتين الكرتين مسحوبتان من الصندوق a الصندوق b الصندوق c

الحل : افرض أن A_1 تمثل حادثة اختيار الصندوق A_2 , a تمثا حادثة اختيار الصندوق A_3 ، b تمثل حادثة اختيار الصندوق c ، وافرض ان A تمثل حادثة اختيار كرتين من صندوق معين بحيث ان احداهما بيضاء والاخرى حمراء . عندئذٍ $P_r(A_1) = P_r(A_2) = P_r(A_3) = \frac{1}{3}$ واذاكان السحب قد تم من الصندوق a عندئذ

$$P_r(A \mid A_1) = \frac{C_1^1 \cdot C_1^3}{C_2^6} = \frac{1}{5}$$

واذا كان السحب قد تم من الصندوق b فان

$$P_r(A \mid A_2) = \frac{C_1^2 \cdot C_1^1}{C_2^4} = \frac{1}{3}$$

اما اذا كان السحب قد تم من الصندوق c فان

$$P_r(A \mid A_3) = \frac{C_1^4 \cdot C_1^3}{C_2^{12}} = \frac{2}{11}$$

عليه فان

$$\sum_{i=1}^{3} P_r(A_i) \cdot P_r(A \mid A_i) = \frac{1}{3} \cdot \frac{1}{5} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{11}$$

$$= \frac{118}{495}$$

وبذلك فان

$$P_r(A_1 \mid A) = \frac{33}{118}, P_r(A_2 \mid A) = \frac{55}{118}, P_r(A_3 \mid A) = \frac{30}{118}$$

واضح من هذا المثال ان

$$\sum_{i=1}^{3} P_r(A_i \mid A) = P_r(\Omega) = 1$$

تمارين

١ ـ ١ : لتكن A_1, A_2 مجموعتين معرفتين في Ω . جد $A_1 \cup A_2$ ، $A_1 \cap A_2$ لكل حالة من الحالات التالية :

أ ـ $A_1 = \{ x : x = 0, 1, 2, 3, 4, 5 \}$, $A_2 = \{ x : x = 2, 4, 6, 8, 10 \}$

ب ـ $A_1 = \{ x : x = -1, -2, -3 \}$, $A_2 = \{ x : x = -1, -2, 0, 1 \}$

جـ ـ $A_1 = \{ x : 2 < x < 6 \}$, $A_2 = \{ x : 4 < x < 10 \}$

د ـ $A_1 = \{ x : 2 < x < 10 \}$, $A_2 = \{ x : x = 2, 4, 6, 8 \}$

١ ـ ٢ : جد متممة المجموعة A ازاء الفضاء Ω لكل مما يلي :

أ ـ $\Omega = \{ x : x = 0, 1, 2, \ldots, n \}$, $A = \{ x : x > 10 \}$

ب ـ $\Omega = \{ x : x = 3, 4, 5, 6, 7, 8 \}$, $A = \{ x : x = 4, 6, 8 \}$

جـ ـ $\Omega = \{ x : x = 0, 1, 2, \ldots \}$, $A = \{ x : x \leq 4 \}$

د ـ $\Omega = \{ x : x \geq 0 \}$, $A = \{ x : 2 < x < 10 \}$

١ ـ ٣ : استخدم مخططات ڤين لتوضيح العمليات التالية : :

$$(A_1 \cup A_2) \cap A_3, (A_1 \cap A_2) \cup A_3, (A_1 \cap A_2) \cup (A_1 \cap A_3)$$
$$(A_1 \cap (A_2 \cup A_3))^c, (A_1 \cup (A_2 \cap A_3))^c$$

١ ـ ٤ : افرض ان

$P_r(x = i) = \frac{1}{15}, i = 1, 2, \ldots 15$ وان $\Omega = \{ x : x = 1, 2, \ldots, 15 \}$

جد احتمال حدوث كل حادثة مما يلي :

x عدد زوجي ، x عدد فردي ، $1 < x \leq 8, 5 \leq x < 14, x \geq 4$

x عدد يقبل القسمة على 3 ، x عدد يقبل القسمة على 3 أو 5 .

x عدد يقبل القسمة على $2, 4, 14 \geq x + 2, 6 \leq x - 1$.

١ ـ ٥ : لمعطيات السؤال (١ ـ ٤) وبفرض ان $A_1 = \{ x : x \leq 8 \}$ ، $A_2 = \{ x : 6 < x \leq 10 \}$, $A_3 = \{ x : 2 \leq x \leq 7 \}$. جد

$$P_r(A_1 \cup A_2 \cup A_3), P_r(A_1 \cap A_2 \cap A_3)$$

١ ــ ٦ : لتكن A_1, A_2, A_3 ثلاث حوادث تنص بان البطارية المسحوبة عشوائيا من وجبة انتاج معينة كانت منتجة من قبل احدى المكائن الثلاث a, b, c. وافرض ان الحادثة A تنص بان هذه البطارية معيبة. ليكن $P_r(A_1) = 0.20$ ، $P_r(A_2) = 0.45$ ، $P_r(A_3) = 0.35$ وان $P_r(A \mid A_1) = 0.05$ ، $P_r(A \mid A_2) = 0.03$ ، $P_r(A \mid A_3) = 0.02$. ماهو احتمال ان هذه البطارية قد انتجت من قبل الماكنة a ، الماكنة b ، الماكنة c.

١ ــ ٧ : عند رمي زهري نرد مرة واحدة ماهو احتمال :

أ ــ ان مجموع النقاط الظاهرة على وجهيهما اكبر من 8 علماً ان عدد النقاط الظاهرة على وجه احدهما هو 6 .

ب ــ ان مجموع النقاط الظاهرة على وجهيهما اقل من 11 علماً ان عدد النقاط الظاهرة على وجه احدهما هو 5 أو 6 .

١ ــ ٨ : اذا علمت ان $m > 0$ ،

$$P_r(A_1) = P, \; P_r(A \mid A_2) = \frac{1}{m}, \; P_r(A \mid A_1) = 1$$

$P_r(A_2) = 1 - P$. برهن ان $P_r(A_1 \mid A) = \dfrac{mP}{1 + (m-1)P}$ ثم جد $P_r(A_2 \mid A)$.

١ ــ ٩ : افرض ان $A_1, A_2, \dots, A_k$ تمثل حوادث مستقلة معرفة في Ω . فاذا كان

$$P_r(A_1^c) \cdot P_r(A_2^c) \dots P_r(A_k^c) = -1 - P_r(A_1 \cup A_2 \cup \dots \cup A_k)$$

وان k ,..., $i = 1, 2$ ، $a > 0$ ، $P_r(A_i) = 1 - \dfrac{1}{a^i}$. برهن أن

$$P_r(\bigcup A_i) = 1 - \frac{1}{a^{\frac{k(k+1)}{2}}}$$

١ ــ ١٠ : اذا كان $P_r(A \mid B) < P_r(A)$ برهن ان $P_r(B \mid A) < P_r(B)$.

١ ــ ١١ : اذا كانت $A, \overline{B}$ حادثتين متنافيتين معرفتين في Ω وان $P_r(A \cup B) > 0$ برهن ان

$$P_r(A \mid A \cup B) = \frac{P_r(A)}{P_r(A) + P_r(B)}$$

١ ــ ١٢ : اذا كانت A, B حادثتين مستقلتين معرفتين في Ω . برهن ان $P_r(A \cup B) = 1 - P_r(A^c) \cdot P_r(B^c)$.

١ ـ ٣ : المتغيرات العشوائية Random variables

ان نظرية الاحتمالات لاتختص بقياس احتمال وقوع حادثة معينة فحسب وانما ايضا في تكوين نموذج رياضي mathematical model يوضح سلوك ظاهرة معينة تتصف بالصفة العشوائية randomness التي تعني انه لادخل للباحث في تحديد اي نتيجة من نتائج التجربة وانما تتحدد وفق الظروف المحيطة بتلك التجربة . فعند رمي زهرنرد فان الرامي لا يعلم مسبقاً بالوجه الذي سيظهر وانما يتحدد ذلك بعد استقرار الزهر ، الا ان الرامي يمتلك معلومات مسبقة عن النتيجة التي ستظهر وهي احد الاعداد (6, 5, 4, 3, 2, 1) ، وبفرض ان الوجه ذات العدد 4 قد ظهرفذلك يمثل نتيجة لهذه التجربة لادخل للرامي في ظهورها او ظهور غيرها . وهذا يعني ان ظهور هذه النتيجة او تلك هو امر محتمل . ان اية حالة مماثلة لهذا المثال تتصف بصفة العشوائية تسمى « تجربة عشوائية random experiment " . وفق هذا المنظور فان الامر يتطلب صياغة وصف رياضي او نموذج رياضي يتصف بطابع احتمالي يفسر سلوك الظاهرة او مجموعة الظواهر قيد التجربة .

١ ـ ٣ ـ ١ : تعريف المتغير العشوائي Definition of random variable

كما سبق ذكره اعلاه فان نتائج اية تجربة عشوائية تتحدد وفق الظروف التجريبية المحيطة بها وان نتائج هذه التجربة سوف لن تكون متشابهة فيما بينها وانما ستكون نتائج مختلفة (لاحظ تجربة رمي زهر النرد) . فاذا رمزنا لاية نتيجة ممكنة للتجربة بالرمز ω وللشيء المطلوب من التجربة بالرمز X (مثلاً في تجربة رمي زهر نرد وملاحظة « عدد النقاط التي ستظهر بعد استقراره » يمكن ان نرمز لهذه العبارة اجمالاً بالرمز X او اي رمز آخر . واذا كنا بصدد دراسة تأثير نوع من الاسمدة في زيادة « انتاجية الدونم الواحد من الحنطة » يمكن ان نرمز لهذه العبارة بالرمز X) . وهذا يعني ان لكل نتيجة من نتائج التجربة ω هنالك عدد حقيقي $X(\omega)$ معرف في فضاء العينة Ω لتلك التجربة . وهذا يعني ان كل عدد حقيقي مثل $X(\omega)$ هو عدد معرف لكل $\omega \in \Omega$. مما تقدم يمكن صياغة التعريف التالي :

افرض ان Ω يمثل فضاء العينة لتجربة عشوائية . ان اية دالة ذات قيمة حقيقية real – valued function معرفة في Ω تأخذ قيماً معرفة في R ، حيث R تمثل

حقل الاعداد الحقيقية ، تسمى متغير عشوائي . وهذا يعني ان المتغير العشوائي X دالة منطلقها Ω domain ومداها R . كذلك يمكن صياغة تعريف آخر للمتغير العشوائي وهو الآتي :

ليكن Ω فضاء العينة لتجربة عشوائية . ان المتغير العشوائي X هو الدالة التي تخصص لكل نتيجة ممكنة مثل $\omega \in \Omega , \omega$ القيمة $X(\omega)$ في المجموعة E حيث $E \subset R$. عليه فان $X(\omega)$ تمثل في الحقيقة عدد حقيقي يعبر عن قيمة المتغير العشوائي X عند نتيجة معينة للتجربة مثل ω . ويمكن التمييز بين نوعين رئيسين من المتغيرات العشوائية التي سترد في الفقرات والفصول اللاحقة هذين النوعين هما :

١ ـ ٣ ـ ٢ : المتغير العشوائي المتقطع (المنفصل)
Discrete random variable

يقال ان X متغير عشوائي متقطع اذا كان فضاء العينة Ω مجموعة قابلة للعد Countable Set سواء كانت مجموعة منتهية ام غير منتهية او بمعنى آخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة متقطع تسمى متغيراً عشوائياً متقطعاً .

مثال (٤) : في تجربة رمي زهر نرد . ان مجموعة القيم الممكنة لهذه التجربة هي المجموعة $\Omega = \{ x : x = 1, 2, 3, 4, 5, 6 \}$ القابلة للعد . فاذن X متغير عشوائي متقطع .

مثال (٥) : افرض ان X يشير الى عدد البطاريات المعيبة في وجبة انتاج مؤلفة من عدد كبير من البطاريات . واضح هنا ان $\Omega = \{ x : x = 0, 1, 2, \ldots \}$ مجموعة قابلة للعد برغم كونها مجموعة غير منتهية . فاذن X متغير عشوائي متقطع .

١ ـ ٣ ـ ٣ : المتغير العشوائي المستمر Continuous random variable

يقال ان X متغير عشوائي مستمر اذا كان فضاء العينة Ω مجموعة غير قابلة للعد سواء كانت منتهية ام غير منتهية ، او بمعنى آخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينه مستمر تسمى متغيراً عشوائياً مستمراً

مثال (٦) : في تجربة لاختيار عدد من الفترة $[0,1]$. واضح ان $\Omega = \{ x : 0 \leq x \leq 1 \}$ مجموعة غير قابلة للعد بسبب وجود عدد غير منته من القيم المعرفة في هذه الفترة . فاذن نستنتج ان X متغير عشوائي مستمر .

مثال (٧) : افرض ان X يشير الى حجم الغازات المنبعثة من انفجار بركاني محتمل الوقوع . واضح هنا ان $\Omega = \{ x : x \geq 0 \}$ مجموعة غير قابلة للعد . فاذن نستنتج ان X متغير عشوائي مستمر .

١ ـ ٣ ـ ٤ : بعض النظريات عن المتغيرات العشوائية .

نستعرض في هذه الفقرة ودون اللجوء الى البرهان بعض النظريات عن المتغيرات العشوائية التي نستفاد منها في فقرات وفصول لاحقة وهي الاتي وبفرض ان X, Y, Z متغيرات عشوائية وان a, b, c ثوابت حقيقية .

١ ـ ان الدوال من الشكل $V_1 = X + Y + Z$, $V_2 = X - Y + Z$, $V_3 = X+Y$.

$V_4 = aX + bY + cZ$, $V_5 = X \cdot Y \cdot Z$, $V_6 = CXY$, $V_7 = \frac{aX}{bY}$, $b \neq 0$

وغيرها هي ايضاً متغيرات عشوائية .

٢ ـ ان $V_1 = \max(X, Y, Z)$, $V_2 = \min(X, Y, Z)$ هي ايضا متغيرات عشوائية .

٣ ـ ان $V_1 = |X + Y + Z|$, $V_2 = |X|$, $V_3 = \frac{1}{|Z|}$ هي متغيرات عشوائية .

٤ ـ ان اية دالة مستمرة بدلالة متغير عشوائي مثل X هي ايضاً متغير عشوائي .

٥ ـ ان اية دالة متزايدة بدلالة متغير عشوائي مثل X هي ايضا متغير عشوائي .

١ ـ ٤ : دوال المتغيرات العشوائية

Functions of random variables

سبق وان ذكرنا في الفقرة (١ ـ ٣) ان نظرية الاحتمالات تختص في صياغة نموذج رياضي لتجربة عشوائية يوضح سلوك ظاهرة معينة (متغير عشوائي) او مجموعة ظواهر معينة (متغيرات عشوائية). وهذا يعني اننا بصدد صياغة دالة تعبر عن سلوك هذا المتغير العشوائي. ورياضياً فان الدالة مثل $f(x)=ax+b, x \varepsilon R$ حيث a, b ثوابت حقيقية، تعبر عن سلوك المتغير X في حقل الاعداد الحقيقية R بحيث أن $f(x)$ تكون معرفة عند أية قيمة من قيم هذا المتغير. كذلك فان $f(x, y)=e^{-(x+y)}, x, y \varepsilon \{x, y): 0 < x, y < \infty\}$ تمثل دالة تعبر عن سلوك المتغيرين X, Y في حقل الاعداد الموجبة بحيث ان $f(x,y)$ تكون معرفة عند أي زوج من القيم (x, y) . ان هذا النوع من الدوال تسمى « دوال نقطة Point functions " بسبب امتلاكها لقيمة حقيقية عند نقطة معرفة في فضاء ذلك المتغير (او تلك المتغيرات). اما من وجهة نظر احتمالية فان يمكن تحديد نوعين رئيسين من دوال المتغيرات العشوائية استناداً الى نوع المتغير العشوائي من حيث كونه متغيراً متقطعاً ام مستمراً. هذان النوعان هما :

١ ـ ٤ ـ ١ : دوال الكتلة الاحتمالية Probability mass functions

وهذه غالبا ماتسمى بالتوزيعات الاحتمالية لمتغيرات عشوائية متقطعة . وعلى فرض ان X متغير عشوائي متقطع لاية نتيجة ممكنة مثل ω معرفة في Ω يوجد عدد حقيقي يقترن بقيمة ω . هذا العدد ماهو الا احتمال الحصول على ω . وهذا يعني وجود دالة تعبر عن قيمة هذا الاحتمال هي $P_r(X = \omega) = P(\omega)$ فاذا كانت $x_1, x_2, \ldots$ تمثل عناصر Ω عندئذ فان كل عنصر من هذه العناصر ستقابله قيمة احتمالية واحدة فقط هي على التوالي $P(x_1), P(x_2), \ldots$ هذه القيمة الاحتمالية تسمى « الكتلة الاحتمالية » المقترنة بالعنصر $x_i, i = 1, 2, \ldots$ بحيث ان $0 \leq P(x_i) \leq 1$ لجميع قيم i . والكتلة الاحتمالية هذه تسمى « دالة الكتلة الاحتمالية » للمتغير العشوائي المتقطع X. وهذا يعني ان دالة الكتلة الاحتمالية دالة منطلقها $E \subseteq \Omega, E$ ، ومداها الفترة $[0, 1]$ كذلك فان المجموعة التي عناصرها هي $P(x_i)$ اي $P=\{P(x_i): 0 \leq P(x_i) \leq 1\}$ تسمى « التوزيع الاحتمالي للمتغير العشوائي Probability distribution of X » .

ان الدالة $P(\omega)$ يجب ان تحقق الشروط التالية كي يسمح لنا ذلك اطلاق تسمية « دالة كتلة احتمالية .p. m. f " عليها وهي :

١ ـ ان $P(x)$ دالة وحيدة القيمة Single – Valued function . اي ان لكل قيمة من قيم X هنالك قيمة واحدة للدالة $P(\omega)$ التي تعبر عن احتمال ان $X = \omega$ اي $P_r(X = \omega)$ . .

٢ ـ ان $P(\omega)$ دالة غير سالبة Non – negative function كونها تعبر عن قيمة احتمالية تقترن بالعنصر ω اي ان... $P(X=\omega_i) \geq 0, i = 1, 2, \ldots$ وهذا يعني ان $P(\omega)$ هي في ذات الوقت متغير عشوائي معرف على الفترة $[0, 1]$ ، وان مخطط الدالة $P(\omega)$ يكون دائماً في الجانب الاعلى من المحور السيني X – axis .

٣ ـ ان مجموع الكتل الاحتمالية المقترنة بعناصرX المعرفة في Ω يجب ان يكون مساوياً للواحد تعبيراً عن احتمال حدوث Ω. اي ان

$$\sum_{x_i \in \Omega} P(x_i) = P_r(\Omega) = 1$$

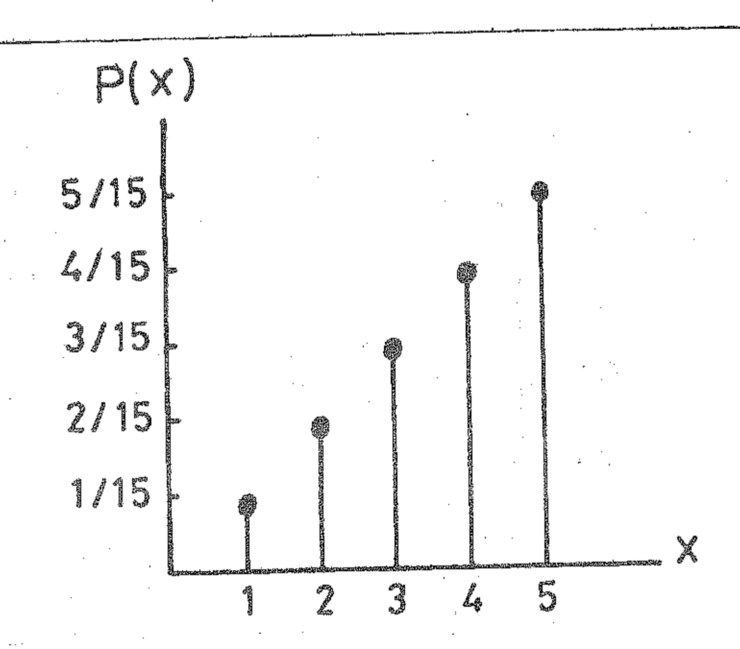
ان معرفتنا المسبقة بدالة الكتلة الاحتمالية الى X ، اي $P(\omega)$ ، تسمح لنا حساب احتمال وقوع اية حادثة (مجموعة جزئية) معرفة في Ω مثل E ، $E \subseteq \Omega$. فمثلاً اذا كانت الحادثة $E = \{\omega : \omega \leq a\}$ حيث a عنصر معرف في Ω ، عندئذ فان احتمال وقوع E هو :

$$P_r(\omega \varepsilon E) = P_r(E) \sum_{\omega \in \Omega} P(x)$$

مثال (٨) : افرض ان X متغير عشوائي يسلك وفق الدالة التالية $P(x) = \frac{x}{15}$ وان $x = 1, 2, 3, 4, 5$ تمثل القيم الممكنة للمتغير X. واضح ان $P(x)$ تمثل دالة كتلة احتمالية كونها دالة وحيدة القيمة ، غير سالبة ، وان مجموع الكتل الاحتماليةالمقترنة بقيم هذا المتغير الممكنة مساوٍ للواحد اي ان :

x :	1	2	3	4	5	
P (x) :	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{5}{15}$	, $\sum_{x=1}^{5} P(x) = 1$

ومخطط الدالة $P(x) = \frac{x}{15}$ هو الموضح بالشكل (١ ـ ٣) :



الشكل (١ ـ ٣) ، مخطط الدالة $P(x) = \frac{x}{15}$

مثال (٩) : افرض ان X متغير عشوائي يسلك وفق الدالة التالية :

$$P(x) = \frac{0.75}{x!(3-x)!}, \quad x = 0, 1, 2, 3$$

بين ان P (x) هي دالة كتلة احتمالية . ثم جد احتمال ان يكون X اقل من او يساوي 2

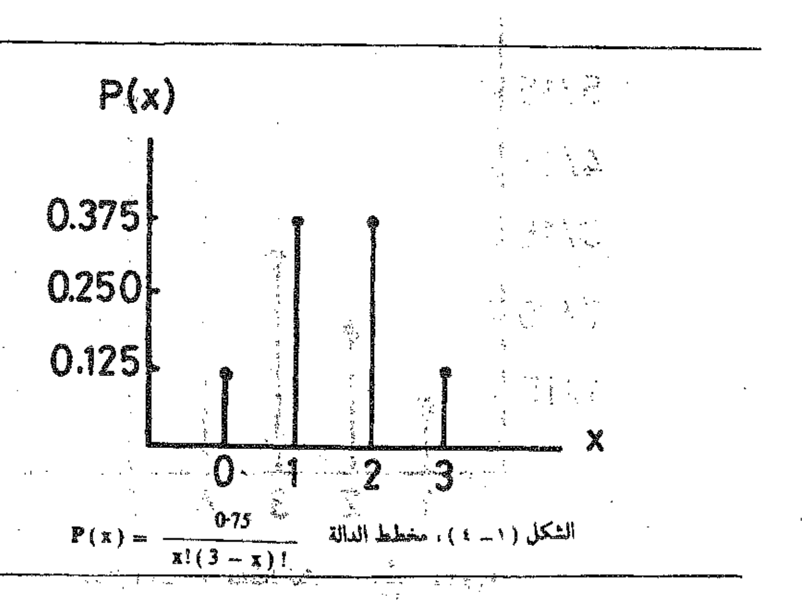
الحل :

x :	0	1	2	3	
P (x) :	0.125	0.375	0.375	0.125	, $\Sigma P(x) = 1$

يتضح من الجدول اعلاه ان لكل قيمة من قيم X هناك قيمة واحدة للدالة P (x) وان قيم P (X) موجبة كذلك فان $\sum_{x=0}^{3} P(x) = 1$ وبتحقيق الشروط الثلاث فذلك يعني ان P (x) هي دالة كتلة احتمالية كذلك فان

$$P_r(X \leq 2) = P(0) + P(1) + P(2) = 0.875$$

والشكل (١ ــ ٤) يوضح مخطط هذه الدالة .



الشكل (١ ــ ٤) ، مخطط الدالة $P(x) = \frac{0.75}{x!(3-x)!}$

مثال (١٠): افرض ان

$$P(x) = \frac{c}{x!(4-x)!}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{4-x}, x = 0, 1, ..., 4$$

تمثل دالة الكتلة الاحتمالية للمتغير العشوائي X . جد قيمة c ثم ارسم مخطط هذه الدالة .

الحل : حيث ان $P(x)$ دالة كتلة احتمالية فذلك يعني ان $\sum_{x=0}^{4} P(x)=1$ فاذن

$$c \sum_{x=0}^{4} \frac{1}{x!(4-x)!}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{4-x} = 1$$

لكن ∵

$$\sum_{x=0}^{4} \frac{1}{x!(4-x)!}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{4-x} = \frac{1}{24}$$

فاذن

$$\frac{c}{24} = 1 \quad \therefore c = 24 = 4!$$

وان مخطط هذه الدالة هو الموضح في الشكل (١ ــ ٥)

الشكل (١ ــ ٥) مخطط الدالة $P(x) = \frac{c}{x!(4-x)!}\left(\frac{1}{4}\right)^{x} \cdot \left(\frac{3}{4}\right)^{4-x}$

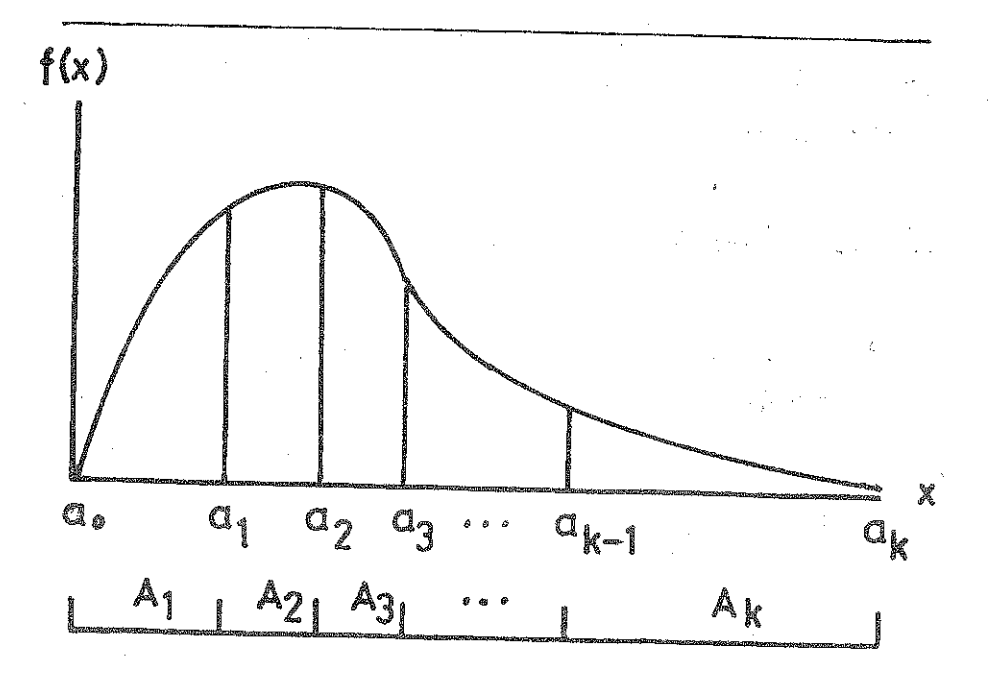
ان التوزيع الاحتمالي المعرف في المثال (٣) يسمى توزيع ثنائي الحدين الدي سيرد ذكره في الفقرة (٥ ـ ٣)

١ ـ ٤ ـ ٢ : دوال الكثافة الاحتمالية Probability density functions

وهذه غالباً ماتسمى بالتوزيعات الاحتمالية لمتغيرات عشوائية مستمرة . وبفرض ان X متغير عشوائي مستمر وان $f(x)$ دالة بدلالة هذا المتغير . ان $f(X = x)$ في هذه الحالة تمثل قيمة الدالة f عند النقطة $X = x$ لكنها لاتعبر عن احتمال ان $X = x$ بسبب ان Ω في هذه الحالة غير قابلة للعد (اي ان قيم المتغير X معرفة في مجموعة الاعداد الحقيقية R) وهذا يعني ان احتمال النقطة $(X = x)$ في حالة المتغيرات العشوائية المستمرة مساوٍ للصفر . وفي هذه الحالة لا يمكن تعريف الدالة الاحتمالية التي تعبر عن سلوك X . لكن يمكن تعريف هذه الدالة ضمن فترة (مجموعة جزيئة) معرفة في Ω مثل $A = \{x : a < x < b\}$ اي حساب احتمال ان $x \in A$ ، اي $P_r(x \in A)$. وحيث ان $P_r(\Omega) = 1$ فان $0 \le P_r(x \in A) \le 1$ طالما ان A مجموعة جزئية من Ω . وهذا يعني ان

$$P_r(x \varepsilon A) = \int_a^b f(x)\,dx$$

اي حساب مقدار كثافة الاحتمال المقترن بالفترة $[a, b]$. وهذا يعني اننا بصدد ايجاد المساحة تحت منحنى الدالة $f(x)$ المحدودة بالفترة $[a, b]$. وبهدف توضيح تعريف هذا النوع من الدوال نفرض ان Ω يمكن تجزئته الى عدد من المجموعات الجزئية مثل $A_1, A_2, \ldots, A_k$ بحيث ان $A_i \cap A_j = \phi, \forall i, j$ وان $A_1 \cup A_2 \cup \cdots \cup A_k = \Omega$. وافرض ان $A_i = \{x : a_i < x < a_{i+1}\}$, $i = 0, 1, \ldots, K$ وهذا يعني ان المساحة تحت منحنى الدالة $f(x)$ سوف تتجزا هي ايضاً الى K من المساحات الجزئية كل منها يقترن باحدى المجموعات الجزئية A_i وكما هو موضح في الشكل (١ ـ ٦) وبفرض ان $\Omega = \{x : a_0 < x < a_k\}$.



الشكل (١ ـ ٦) ، تجزئة الفضاء الى عدد من المجموعات الجزئية

فان كان $P_r(\Omega) = 1$ فذلك يعني ان

$$P_r(\Omega) = \sum_{i=1}^{k} P_r(A_i) = \sum_{i=0}^{k-1} \int_{a_i}^{a_{i+1}} f(x)\,dx = 1$$

وعندئذ يقال ان $f(x)$ هي دالة الكثافة الاحتمالية للمتغير العشوائي X . مما تقدم يمكن القول ان $f(x)$ دالة منطلقها مجموعة المجموعات الجزئية $\{A_i\}$ ومداها الفترة $[0, 1]$. كذلك فان المجموعة التي عناصرها هي $P_r(A_i)$ اي $P = \{P_r(A_i) : 0 \leq P_r(A_i) \leq 1\}$ تسمى التوزيع الاحتمالي الى X . ووفق ماتقدم فان الدالة $f(x)$ يجب ان تحقق الشروط التالية التي تسمح لنا اطلاق تسمية دالة كثافة احتمالية عليها وهي :

١ ـ ان $f(x)$ دالة وحيدة القيمة Single – Valued function اي ان لكل $x \in \Omega$ هنالك قيمة واحدة فقط الى $f(x)$.

٢ ـ ان $f(x)$ دالة غير سالبة لجميع قيم $x \in \Omega$. اي ان $f(x)$ دالة موجبة دائماً . وهذا يعني ان مخطط الدالة $f(x)$ يكون دائماً في الجانب الاعلى من المحور السيني . ان منحنى الدالة $f(x)$ غالباً ما يسمى « المنحنى الاحتمالي Probability Curve للدالة $f(x)$.

٣ ـ ان $\int_{\Omega} f(x)\,dx = 1$ اي ان الاحتمال المقترن بفضاء X يجب ان يكون مساوياً للواحد .

ان معرفتنا المسبقة بدالة الكثافة الاحتمالية للمتغير X تسمح لنا حساب احتمال حدوث اية حادثة معرفة في Ω فمثلاً اذا كانت $A=\{x : x \leq a\}$ حيث a قيمة معرفة في Ω فان احتمال حدوث A هو .

$$P_r(A) = \int_{x \in A} f(x)\,dx \leq 1$$

مثال (١١) افرض ان X متغير عشوائي يسلك وفق الدالة التالية $f(x)=3x^2 , 0 < x < 1$ واضح ان هذه الدالة هي دالة كثافة احتمالية كونها دالة وحيدة القيمة . موجبة وان

$$\int_{\Omega} f(x)\,dx = \int_0^1 3x^2 dx = 1$$

كذلك فان اية تجزئة لفضاء X تقودنا الى تعريف التوزيع الاحتمالي لهذا المتغير فمثلاً لو تم تجزئة Ω الى أربع مجموعات جزئية غير مشتركة من الشكل :

$$A_1 = \left\{ x : 0 \leq x \leq \frac{1}{4} \right\}, A_2 = \left\{ x : \frac{1}{4} < x \leq \frac{1}{3} \right\},$$

$$A_3 = \left\{ x : \frac{1}{3} < x \leq \frac{3}{4} \right\}, A_4 = \left\{ x : \frac{3}{4} < x \leq 1 \right\}$$

عندئذٍ فان :

$$P_r(A_1) = \int_0^{\frac{1}{4}} 3x^2 dx = \frac{1}{64} , P_r(A_2) = \int_{\frac{1}{4}}^{\frac{1}{3}} 3x^2 dx = \frac{37}{1728} ,$$

$$P_r(A_3) = \int_{\frac{1}{3}}^{\frac{3}{4}} 3x^2dx = \frac{665}{1728}, \quad P_r(A_4) = \int_{\frac{3}{4}}^{1} 3x^2dx = \frac{37}{64}$$

وان

$$\sum_{i=1}^{4} P_r(A_i) = 1$$

ان مخطط الدالة f(x) موضح في الشكل (١ ـ ٧) :

الشكل (١ ـ ٧) ، مخطط الدالة $F(x) = 3x^2$

مثال (١٢): افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = ce^{-3x}, x \geq 0$
جد قيمة c . واذا كانت $A = \{x : 1 < x < 3\}, B = \{x : 0 < x < 2\}$
جد $P_r(A), P_r(B), P_r(A \cap B)$. ثم ارسم مخطط الدالة $f(x)$.

الحل : حيث ان f(x) دالة كثافة احتمالية فذلك يعني ان

$$\int_{\Omega} f(x)\,dx = c\int_0^{\infty} e^{-3x}dx = 1 \qquad \therefore c = 3$$

فاذن $f(x) = 3e^{-3x}, x \geq 0$ وبذلك فان

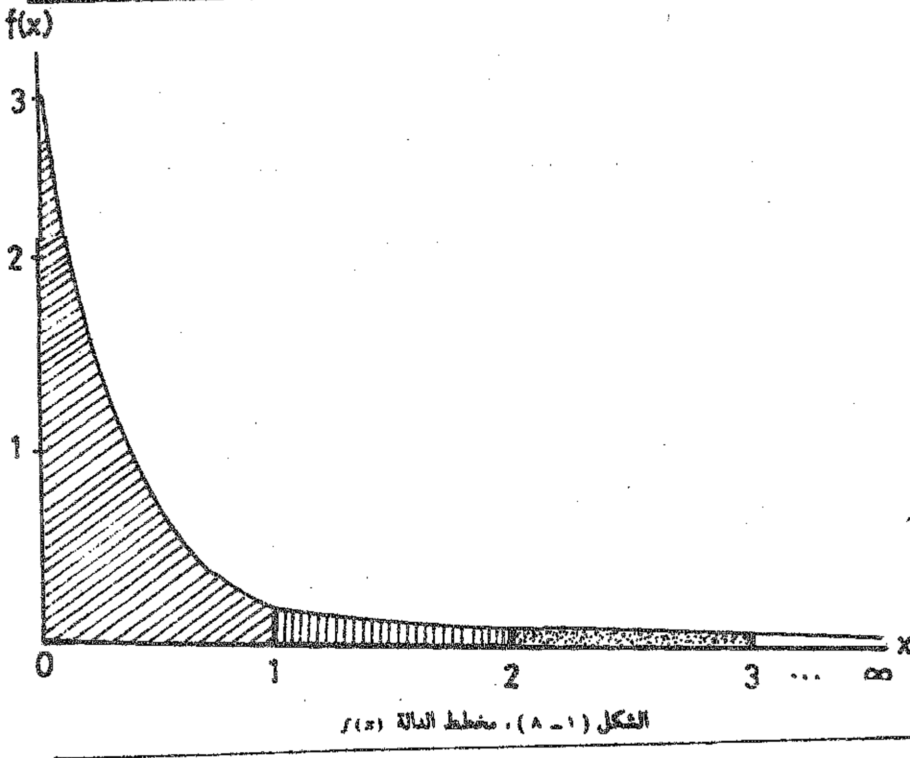
$$P_r(A) = \int_1^3 3e^{-3x}dx = e^{-3} - e^{-9} = 0{\cdot}0496635$$

$$P_r(B) = \int_0^2 3e^{-3x}dx = 1 - e^{-6} = 0{\cdot}9975213$$

وان

$$P_r(A \cap B) = \int_1^2 3e^{-3x}dx = e^{-3} - e^{-6} = 0{\cdot}0473082$$

ان مخطط الدالة f(x) موضح في الشكل (١ ــ ٨) .



الشكل (١ ــ ٨) ، مخطط الدالة $f(x)$

١ ـ ٥ : دالة التوزيع التراكمية

Cumulative distribution function

وتسمى في بعض الاحيان « دالة التوزيع » أو « الدالة التوزيعية » او « دالة التراكم الاحتمالي » . وتعرف هذه الدالة بانها قيمة الاحتمال المتراكم لغاية قيمة معطاة من قيم المتغير العشوائي X المعرفة في Ω . وغالباً مايرمز لهذه الدالة بالشكل $F(x)$. وهذا يعني ان $F(x) = P_r(X \leq x)$ نستشف مما تقدم ان $F(x)$ دالة غير متناقصة non – decreasing function تعبيراً عن عملية تراكم احتمالي . ان لهذه الدوال اهمية كبيرة في حساب مايسمى « القيم الجدولية » او « القيم الحرجة critical values » للتوزيعات الاحتمالية التي سيرد ذكرها في فصول لاحقة .

١ ـ ٥ ـ ١ : دالة التوزيع للمتغيرات المتقطعة:

افرض ان X متغير عشوائي متقطع بدالة كتلة احتمالية $P(x)$ معرفة قيمة في Ω وافرض ان x قيمة من قيم X المعرفة في Ω . عندئذٍ فان :

$$F(x) = P_r(X \leq x) = \sum_{k \leq x} P(X = k)$$

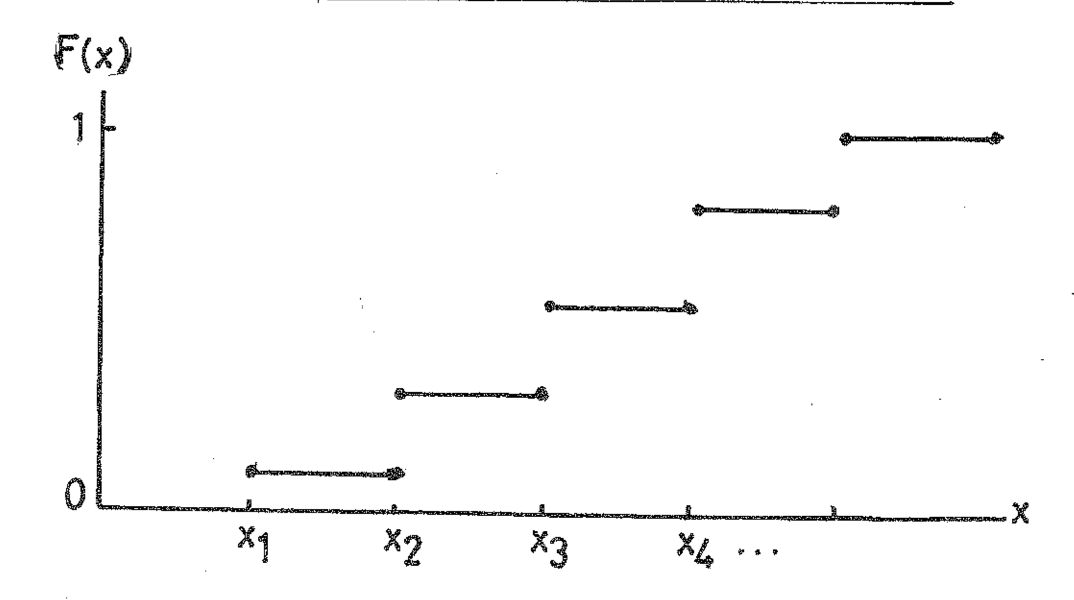
كذلك فان

$$\lim_{x \to -\infty} F(x) = F(-\infty) = 0, \quad \lim_{x \to \infty} F(x) = F(\infty) = 1$$

وهذا يعني ان $0 \leq F(x) \leq 1$. عليه يمكن القول ان $F(x)$ متغير عشوائي معرف على الفترة $[0, 1]$. واذا كانت $F(x)$ معلومة عندئذٍ يمكن حساب قيمة الكتلة الاحتمالية المقترنة بالعنصر x_i وفق مايلي :

ان $F(x_i)$ يمثل التراكم لغاية x_i وان $F(x_{i-1})$ يمثل التراكم الاحتمالي لغاية x_{i-1} . عليه فان $P(x_i) = F(x_i) - F(x_{i-1})$

ان مخطط الدالة $F(x)$ يكون على شكل متدرج وصولا الى قيمة $F(x)$ المساوية الى واحد ، وكما هو موضح في الشكل (١ ـ ٩) :



شكل (١ ــ ٩) ، مخطط الدالة F (x)

مثال (١٣) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{1}{4}\left(\frac{3}{4}\right)^x$ ، $x = 0, 1, 2, \ldots$. جد الدالة التوزيعية .

الحل :

$$F(x) = P_r(X \leq x) = \sum_{k=0}^{x} \left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^k = \frac{1}{4} \sum_{k=0}^{x} \left(\frac{3}{4}\right)^k$$

لكن $\sum_{k=0}^{x}\left(\frac{3}{4}\right)^k$ يعبر عن مجموع حدود متوالية هندسية نهائية حدها الاول مساوٍ للواحد واساسها مساوٍ الى $\frac{3}{4}$. وهذا يعني ان

$$\sum_{k=0}^{x} \left(\frac{3}{4}\right)^k = \frac{1-(3/4)^{x+1}}{1-(3/4)} = 4\left(1-(3/4)^{x+1}\right)$$

فاذن

$$F(x) = 1 - \left(\frac{3}{4}\right)^{x+1}, \quad x = 0, 1, 2, \ldots$$

ومن خلال هذه الدالة يمكن حساب الاحتمال المتراكم لغاية اية قيمة من قيم X المعرفة في Ω بمجرد التعويض عن تلك القيمة في الدالة F (x) . فمثلاً

$$F(0) = 1 - \frac{3}{4} = 0{\cdot}25\,, F(2) = 1 - \left(\frac{3}{4}\right)^3 = 0{\cdot}578125\,,$$

$$F(3) = 1 - \left(\frac{3}{4}\right)^4 = 0{\cdot}6835938, F(10) = 1 - \left(\frac{3}{4}\right)^{11}$$

$$= 0{\cdot}9577649$$

كذلك فان

$$.\,P(X = 3) = F(3) - F(2) = 0{\cdot}1054688$$

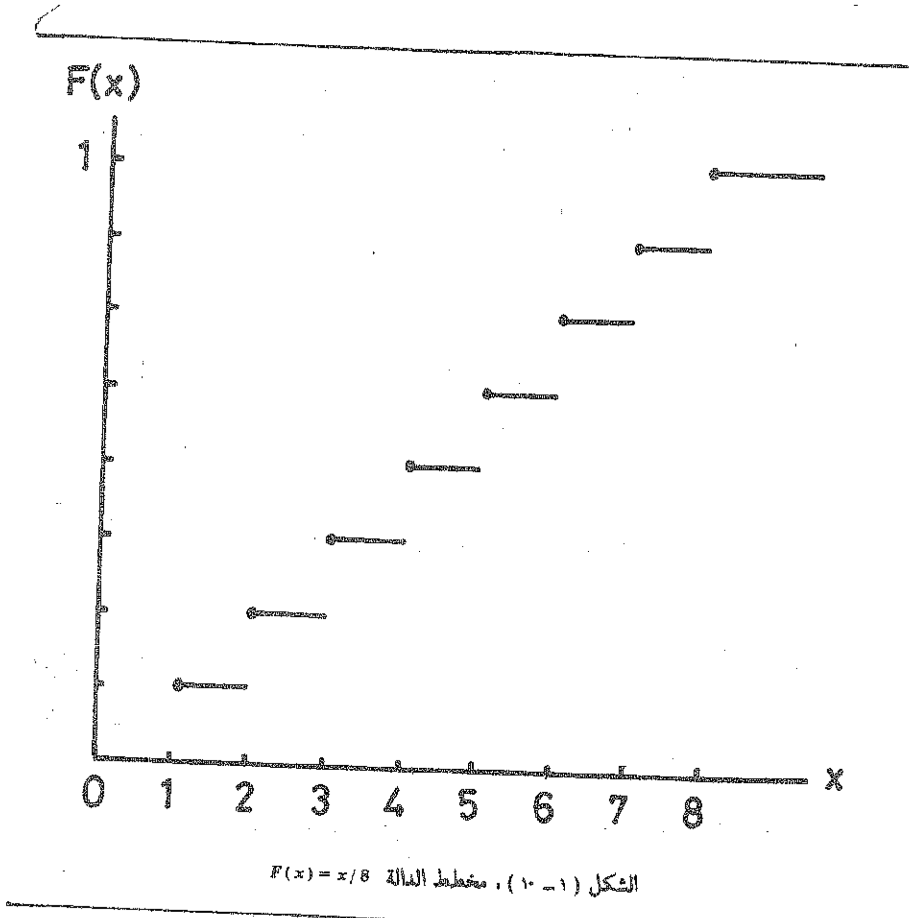
مثال (١٤) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{1}{8}$ $x = 1, 2, \ldots, 8$ جد الدالة التوزيعية ثم ارسم هذه الدالة .

الحل : $F(x) = P_r(X \leq x) = \sum_{k=1}^{x} \frac{1}{8} = \frac{x}{8}\,, x = 1, 2, \ldots, 8$

واضح ان

$$\begin{aligned} F(x) &= 0 &, x < 1 \\ &= 1/8 &, x \leq 1 \\ &= 2/8 &, x \leq 2 \\ &= 3/8 &, x \leq 3 \\ &= 4/8 &, x \leq 4 \\ &= 5/8 &, x \leq 5 \\ &= 6/8 &, x \leq 6 \\ &= 7/8 &, x \leq 7 \\ &= 1 &, x \leq 8 \end{aligned}$$

والشكل (١ ـ ١٠) يوضح مخطط هذه الدالة .



الشكل (١ ــ ١٠) ، مخطط الدالة $F(x) = x/8$

١ ــ ٥ ــ ٢ : دالة التوزيع للمتغيرات المستمرة

افرض ان X متغير عشوائي مستمر بدالة كثافة احتمالية $f(x)$ وان $\Omega = \{x : -\infty < x < \infty\}$. لتكن x قيمة من قيم X المعرفة في Ω . عندئذ فان دالة التوزيع هي :

$$F(x) = P_r(X \leq x) = \int_{-\infty}^{x} f(\omega)\, d\omega$$

وفيما يلي بعض الملاحظات عن هذه الدالة ،

١ ــ ان

$$\lim_{x \to -\infty} F(x) = F(-\infty) = \int_{-\infty}^{-\infty} f(\omega)\, d\omega = 0$$

$$\lim_{x \to \infty} F(x) = F(\infty) = \int_{-\infty}^{\infty} f(\omega)\, d\omega = P_r(\Omega) = 1$$

٢ ــ يمكن التعبير عن احتمال وقوع X في فترة معينة مثل [a, b] معرفة في Ω بدلالة F(x) وكما يلي :

$$P_r(a < X < b) = \int_a^b f(x)\, dx = \int_{-\infty}^{b} f(x)\, dx - \int_{-\infty}^{a} f(x)\, dx$$

$$= F(b) - F(a)$$

٣ ــ حيث ان F(x) هي نتيجة لتكامل الدالة f(x) فذلك يعني ان مشتقة F(x) نسبة الى X ماهي الا f(x) . اي ان .

$$\frac{dF(x)}{dx} = f(x) \rightarrow dF(x) = f(x)\, dx$$

وهذا غالباً مايسمى « التفاضل الاحتمالي » للمتغير X . ان هذه الملاحظة مهمة جداً في موضوع استنتاج دالة الكثافة الاحتمالية لمتغير عشوائي مثل X علمت دالته التوزيعيه . وسوف نستعرض هذا الموضوع وبشكل مفصل في الفقرة (٧ ــ ٢) .

مبرهنة :

ان F(x) دالة مستمرة نحو الجانب الايمن وغير متناقصة ، وهي في ذات الوقت متغير عشوائي مستمر معرف على الفترة [0,1] . فاذا كان $Y = F(X)$ متغير مستمر فان $g(y) = 1, 0 \leq y \leq 1$

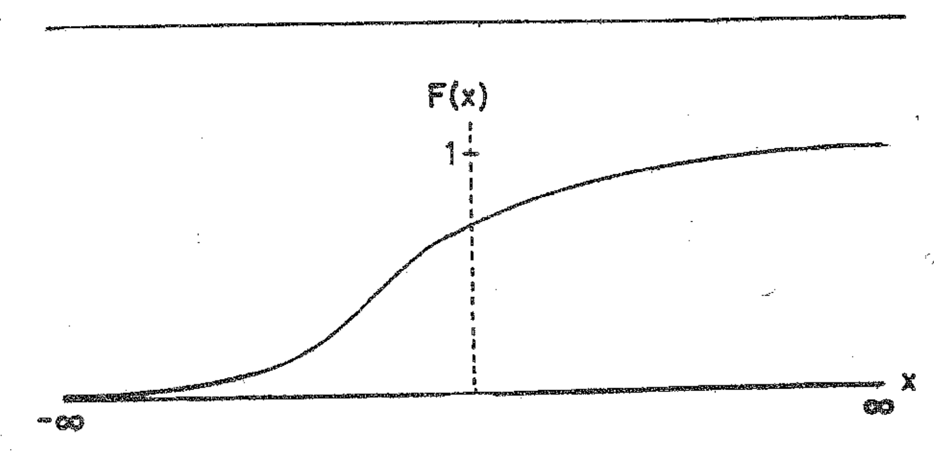
البرهان : افرض ان الدالة التوزيعية الى y هي $G(y)$ ، وذلك يعني ان

$$G(y) = P_r(Y \leq y) = P_r(F(x) \leq y)$$
$$= P_r(x \leq F^{-1}(y)) = F(F^{-1}(y))$$
$$\therefore\ G(y) = F(F^{-1}(y)) = y$$

وبتفاضل الطرفين نسبة الى y نحصل على

$$g(y) = G'(y) = 1\ ,\ 0 \leq y \leq 1$$

ان مخطط الدالة $F(x)$ بشكل عام هو الموضح في الشكل (١ ــ ١١) :



الشكل (١ ــ ١١) : مخطط الدالة $F(x)$

مثال (١٥) : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = e^{-x}, x \geq 0$ جد $F(x)$ ثم احسب $P_r(1 < x < 3)$ مع رسم الدالة

الحل :

$$F(x) = \int_{-\infty}^{x} f(\omega)\,d\omega = \int_{0}^{x} e^{-\omega}d\omega = 1 - e^{-x}, x \geq 0$$

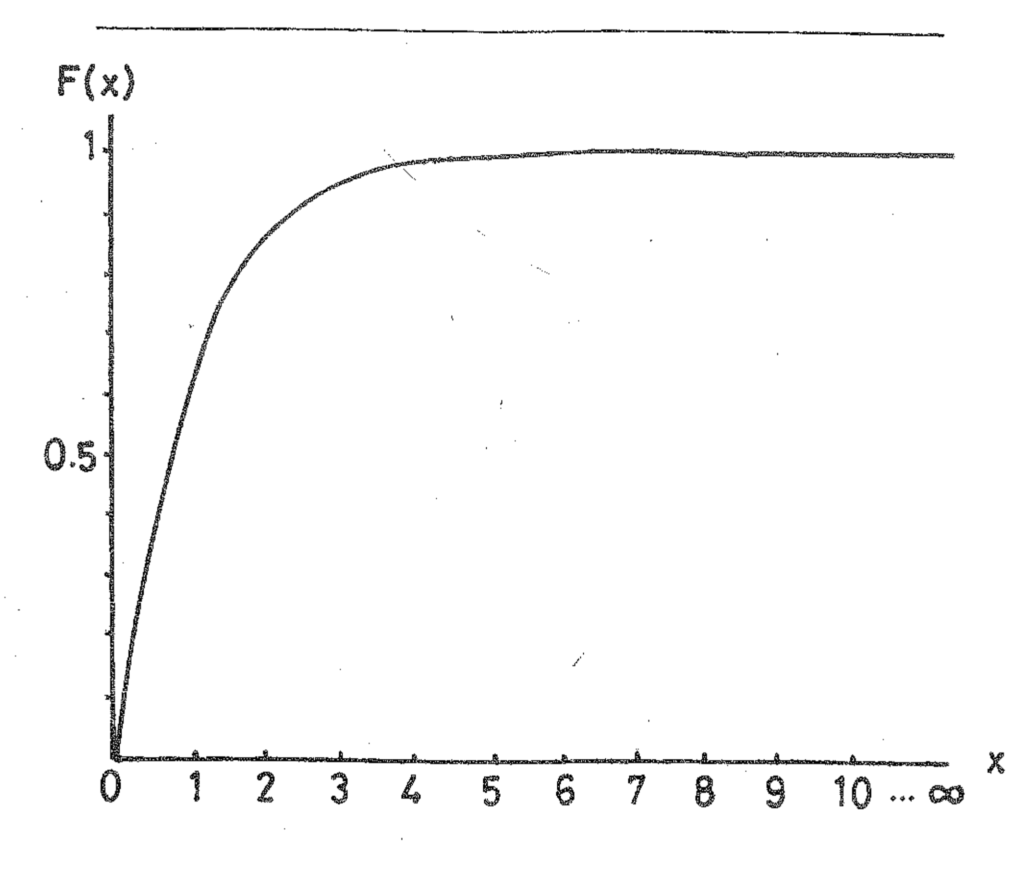
لاحظ من هذه الدالة ان

$$F(0) = 0, F(\infty) = 1, \frac{dF(x)}{dx} e^{-x} = f(x)$$

كذلك فان

$$P_r(1 < x < 3) = F(3) - F(1) = (1 - e^{-3}) - (1 - e^{-1}) = 0{\cdot}3180924$$

والشكل (١ ــ ١٢) يوضح مخطط هذه الدالة .



الشكل (١ ــ ١٢) ، مخطط الدالة $F(x) = 1 - e^{-x}$

مثال (١٦) : اذا علمت ان الدالة التوزيعية لمتغير عشوائي مستمر مثل X هي :

$$F(x) = 0 \quad , x < -3$$

$$= \frac{1}{6}(x+3) \quad , -3 \leq x \leq 3$$

$$= 1 \quad , x \leq 3$$

جد دالة الكثافة الاحتمالية الى X .

الحل :

$$f(x) = F'(x) = \frac{1}{6} \quad , -3 \leq x \leq 3$$

$$= 0 \quad , \text{other wise}$$

مثال (١٧) : افرض ان X يمثل عمر نوع من الصمامات الالكترونية (مقاس بالساعات) بدالة كثافة احتمالية $x \geq 200$, $f(x) = \frac{200}{x^2}$. جد الدالة التوزيعية . ماهو احتمال عمر صمام معين اقل من 300 ساعة .

الحل :

$$F(x) = \int_{200}^{x} \frac{200}{\omega^2}\, d\omega = 1 - \frac{200}{x} \quad , x \geq 200$$

$$P_r(X < 300) = F(300) = 1 - \frac{200}{300} = 0.333$$

تمارين

١ ـ ١٣ : جد قيمة الثابت C لكل حالة من الحالات التالية بحيث ان $f(x)$ هي دالة كثافة احتمالية وان $P(x)$ هي دالة كتلة احتمالية :

$$f(x) = cxe^{-x},\ x > 0 \qquad ,\ P(x) = c\left(\frac{4}{5}\right)^{x},\ x = 1, 2, \ldots$$

$$f(x) = ce^{-x} \quad ,\ a < x < b\ ,\ P(x) = \frac{c4^{x}}{x!} \qquad ,\ x = 0, 1, 2, \ldots$$

$$f(x) = e^{-cx} \quad ,\ x \geq 0 \qquad ,\ P(x) = \frac{x}{c} \qquad ,\ x = 1, 2, 3, 4.$$

١ ـ ١٤ : افرض ان X متغير عشوائي بدالة كتلة احتمالية

$$P(x) = \frac{6!}{x!(6-x)!}\left(\frac{1}{3}\right)^{x}\cdot\left(\frac{2}{3}\right)^{6-x},\ x = 0, 1, \ldots, 6$$

جد مايلي :

أ ـ التوزيع الاحتمالي الى X

ب ـ $P_r(1 < X \leq 5), P_r(X > 2), P_r(X \leq 3)$

جـ ـ ارسم مخطط الدالة $P(x)$

١ ـ ١٥ : افرض ان X متغير عشوائي بدالة كثافة احتمالية

$$f(x) = \frac{1}{x^{2}},\ 1 < x < \infty$$

يطلب اجراء مايلي :

أ ـ ايجاد الدالة التوزيعية الى X مع رسم مخطط هذه الدالة .

ب ـ اذا كانت $A = \{x: 3< x< 9\}$, $B = \{x: 5<x <12\}$. جد

$P_r(A \cap B), P_r(A \cup B), P_r(B)\ \ P_r(A)$

١ ـ ١٦ : افرض ان $f(x)=e^{-x}, x \geq 0$ تمثل دالة الكثافة الاحتمالية الى X . وافرض ان $Y = F(X)$ بين ان دالة الكثافة الاحتمالية للمتغير Y هي $f(y) = 1, 0 \leq y \leq 1$

١ ـ ١٧ : افرض ان دالة الكتلة الاحتمالية لتغير عشوائي X مرصوفة بالتوزيع الاحتمالي التالي :

x :	-4	-3	-2	-1	0	1	2	3	4
$P(x)$:	P	$2P$	$2P$	$3P$	$3P^2$	$4P^2$	$5P^2$	$2P$	P

يطلب اجراء مايلي :

أ ـ جد قيمة P ، ثم ارسم هذه الدالة .

ب ـ جد الدالة التوزيعية ثم ارسم هذه الدالة .

جـ ـ جد $P_r(-1 < X \leq 3), P_r(X \geq -2), P_r(X \leq 1)$

١ ـ ١٨ : ليكن X متغير عشوائي يتوزع وفق الدالة $f(x) = \sin x, 0 \leq x \leq \frac{\pi}{2}$

أ ـ بين ان $f(x)$ هي دالة كثافة احتمالية ثم ارسم مخطط هذه الدالة .

ب ـ جد الدالة التوزيعية ثم ارسم مخطط هذه الدالة .

جـ ـ جد $P_r\left(\frac{\pi}{5} < X < \frac{\pi}{3}\right), P_r\left(X > \frac{\pi}{4}\right), P_r\left(X < \frac{\pi}{6}\right)$

١ ـ ١٩ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = k\sec^2 x$ ، $0 \leq x \leq \frac{\pi}{6}$ وان k ثابت حقيقي . يطلب اجراء مايلي :

أ ـ جد قيمة k ثم ارسم مخطط الدالة $f(x)$

ب ـ جد $F(x)$ ثم ارسم مخطط هذه الدالة .

جـ ـ اذا كانت $B = \left\{x: \frac{\pi}{4} < x < \frac{\pi}{2}\right\}, A = \left\{x: \frac{\pi}{5} < x < \frac{\pi}{3}\right\}$

جد $P_r(A \cap B), P_r(A \cup B), P_r(B), P_r(A)$.

١ ــ ٢٠ : ليكن X متغيراً عشوائياً بدالة كثافة احتمالية $f(x)=\frac{k}{1+x^2},-\infty<x<\infty$
يطلب اجراء مايلي :

أ ــ جد قيمة الثابت k ثم ارسم مخطط الدالة $f(x)$

ب ــ بين ان $F(x) = \frac{1}{\pi}\left(\tan^{-1}x + \frac{\pi}{2}\right)$ ثم ارسم مخطط هذه الدالة .

جـ ــ اذا كانت $A = \{x : x \leq 5\}$, $B = \{x : x \leq 8\}$ جد
$P_r(A)$, $P_r(B)$, $P_r(A \cap B)$.

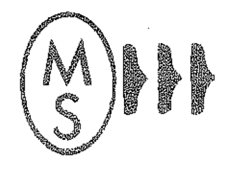
١ ــ ٢١ : لوحظ في مدرج احد المطارات ان عدد الدقائق التي تنتظرها الطائرة لحين مجيء دورها للاقلاع هو متغير عشوائي بدالة توزيعية هي :

$$F(x) = 1 - e^{-0.3x} \quad , x \geq 0$$
$$= 0 \quad , x < 0$$

يطلب اجراء مايلي :
أ ــ جد دالة الكثافة الاحتمالية الى X ثم ارسم مخطط هذه الدالة .
ب ــ ماهو احتمال ان تنتظر طائرة معينة اكثر من عشرة دقائق ؟

١ ــ ٢٢ : صندوق يحتوي على 15 كرة ست منها بيضاء والبقية سوداء . اختيرت من هذا الصندوق خمس كرات عشوائياً . وبفرض ان X يمثل عدد الكرات البيضاء الموجودة ضمن العينة المختارة . ماهي الدالة الاحتمالية الى X ؟ وماهو احتمال الحصول على الاقل ثلاث كرات بيضاء ؟

١ ــ ٢٣ : لوحظ من خلال الخبرة السابقة في احد مصانع انتاج البطاريات الجافة ان نسبة عدد البطاريات المعيبة هي 5 % . اختيرت عينة مؤلفة من 10 بطاريات من انتاج احدى الوجبات وبفرض ان X يمثل عدد البطاريات المعيبة في هذه العينة . ماهي الدالة الاحتمالية الى X ؟ ماهو احتمال وجود على الاكثر بطارية واحدة معيبة في هذه العينة ؟ ماهو احتمال وجود على الاقل ثلاث بطاريات غير معيبة ؟ .





التوقع الرياضي والدوال المولدة للعزوم

الفصل الثاني

التوقع الرياضي والدوال المولدة للعزوم

نستعرض في هذا الفصل مفهومين اساسين في النظرية الاحصائية هما التوقع الرياضي والدوال المولدة للعزوم . ونظراً لاهميتها فقد ارتأيت تخصيص هذا الفصل لدراستهما بشكل مفصل بسبب اعتماد الكثير من الفقرات اللاحقة عليهما .

٢ ـ ١ : التوقع الرياضي Mathematical Expectation

افرض ان X متغير عشوائي بدالة كتلة احتمالية P(x) او كثافة احتمالية f(x) وان Ω يمثل فضاء العينة الى X . لتكن g(x) دالة بدلالة X . ان الدالة g(x) في الحقيقة هي الاخرى متغير عشوائي بسبب اعتمادها على X . ويعرف التوقع الرياضي للدالة g(x) بانه عملية ايجاد متوسط g(x) ويرمز لهذه العملية بالشكل E[g(x)] الذي غالباً مايسمى « التوقع الرياضي للدالة g(x) » او « القيمة المتوقعة للدالة g(x) » او « متوسط الدالة y(x) » . فعلى سبيل المثال في تجربة رمي زهر نرد وبفرض ان X يمثل عدد النقاط الظاهرة على وجه الزهر بعد استقراره وبفرض ان g(x) = x عندئذٍ E[g(x)]=EX تعني في الحقيقة عملية حساب « متوسط عدد النقاط » في هذه التجربة التي نتائجها الممكنة هي { 1, 2, 3, 4, 5, 6 } اي ان :

$$EX = \frac{1}{6} \sum_{x=1}^{6} x = \frac{21}{6}$$

لاحظ ان EX يمثل موقع « مركز القيم » المعرفة في Ω على المحور السيني . واذا كانت $g(x)=x^2$ فان EX^2 تعني عملية حساب « متوسط مربعات عدد النقاط » في هذه التجربة ، اي ان :

$$EX^2 = \frac{1}{6} \sum_{x=1}^{6} x^2 = \frac{91}{6}$$

٢ ـ ١ ـ ١ : التوقع الرياضي في حالة المتغيرات المتقطعة :

في حالة المتغيرات المتقطعة وبفرض ان $P(x)$ تمثل دالة الكتلة الاحتمالية الى X المعرفة قيمه في Ω فان القيمة المتوقعة للدالة $g(x)$ يمكن حسابها على النحو الآتي :

$$E[g(x)] = \sum_{x\in\Omega} g(x) . P(X = x)$$

بشرط ان هذا المجموع متقارب على نحو مطلق absolutely convergent اي مانعنيه ان

$$\sum_{x\in\Omega} |g(x)\cdot P(X = x)| = \sum_{x\in\Omega} |g(x)| \cdot P(X = x) < \infty$$

ان هذا الشرط يعني ان $E[g(x)]$ معرف . اما اذا كان المجموع متباعداً diverge عندئذٍ يقال ان $E[g(x)]$ غير معرف .

مثال (١) : افرض ان $P(x) = \frac{e^{-1}}{x!}, x = 0, 1, 2, \ldots$ وان $g(x) = x!$ عندئذٍ

$$E[g(x)] = E(X!) = \sum_{x=0}^{\infty} x! . \frac{e^{-1}}{x!} = e^{-1} \sum_{x=0}^{\infty} 1$$

واضح في هذه الحالة ان $\sum_{x=0}^{\infty} 1$ متباعد . عليه فان $E(X!)$ غير معرف . في حين اذا كانت $g(x) = x(x-1)$ عندئذٍ :

$$E[g(x)] = EX(X-1) = \sum_{x=0}^{\infty} x(x-1) . \frac{e^{-1}}{x!}$$

$$= \sum_{x=2}^{\infty} \frac{e^{-1}}{(x-2)!} = \sum_{y=0}^{\infty} \frac{e^{-1}}{y!} = e^{-1} \sum_{y=0}^{\infty} \frac{1}{y!} \quad , y = x - 2$$

لاحظ في هذه الحالة ان المجموع متقارب نحو العدد e . وهذا يعني ان التوقع اعلاه موجود ومساوٍ الى $e^{-1} . e = 1$.

مثال (٢) : افرض ان $p(x) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^x, x = 0, 1, \ldots$ وان $g(x) = 6^x$ ان توقع الدالة $g(x)$ هو :

$$E[g(x)] = E6^x = \sum_{x=0}^{\infty} 6^x . \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^x$$

$$= \frac{3}{4}\sum_{x=0}^{\infty}\left(\frac{3}{2}\right)^x$$

لاحظ ان المجموع متباعد . وهذا يعني ان $E6^x$ غير معرف . في حين اذا كانت $g(x) = 2^x$ عندئذٍ .

$$E2^x = \sum_{x=0}^{\infty} 2^x . \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^x = \frac{3}{4}\sum_{x=0}^{\infty}\left(\frac{1}{2}\right)^x$$

لاحظ ان المجموع متقارب نحو العدد $2 = \frac{1}{1 - 0.5}$ وهذا يعني ان $E2^x$ موجود ومساوٍ الى $\frac{3}{4} . 2 = \frac{3}{2}$

واضح مما تقدم ان مشكلة عدم امكانية ايجاد $E[g(x)]$ في بعض الاحيان تبرز بسبب كون ان Ω مجموع غير منتهية . في حين اذا كانت Ω مجموعة منتهية فان $E[g(x)]$ عادة يمكن ايجاده.

٢ ــ ١ ــ ٢ : التوقع الرياضي في حالة المتغيرات المستمرة .

في حالة المتغيرات المستمرة وبفرض ان $f(x)$ تمثل دالة كثافة احتمالية الى X المعرفة قيمه في Ω فان القيمة المتوقعة للدالة $g(x)$ يمكن حسابها على النحو الاتي :

$$E[g(x)] = \int_{\Omega} g(x)f(x)\,dx$$

بشرط ان التكامل متقارب على نحو مطلق . اي ان .

$$\int_{\Omega} | g(x) f(x) | dx = \int_{\Omega} | g(x) | f(x) dx < \infty$$

اما اذا كان التكامل متباعداً عندئذ يقال ان توقع الدالة $g(x)$ غير معرف .

مثال (٣) : افرض ان $f(x) = e^{-x} \quad x > 0$ وان $g(x) = e^{x}$ ان توقع $g(x)$ هو

$$E[g(x)] = Ee^{X} = \int_{0}^{\infty} e^{x} . e^{-x} dx = \int_{0}^{\infty} dx = \lim_{x \to \infty} x$$

لاحظ ان التكامل متباعد وهذا يعني ان Ee^{X} غير معرف . في حين اذا كانت $g(x) = 3x$ فان

$$E(3X) = \int_{0}^{\infty} 3x . e^{-x} dx = 3 \int_{0}^{\infty} x e^{-x} dx$$

وباستخدام التكامل بطريقة التجزئة يمكن البيان ان التكامل متقارب ومساوِ الى واحد .فاذن $E(3X)$ موجود ومساوِ الى 3 .

مثال (٤) : لتكن $f(x) = \frac{1}{x^2} \; , x > 1$ تمثل دالة الكثافة الاحتمالية الى X وافرض ان $g(x) = bx^2$ ان توقع الدالة $g(x)$ هو :

$$E(bX^2) = \int_{1}^{\infty} bx^2 . \frac{1}{x^2} dx = b \int_{1}^{\infty} dx = b \left(\lim_{x \to \infty} x - 1 \right)$$

لاحظ ان التكامل متباعد وهذا يعني ان $E(bX^2)$ غير موجود . واذا كانت $g(x) = \sqrt{X}$ فان :

$$E(\sqrt{X}) = \int_1^\infty x^{\frac{1}{2}} \frac{1}{x^2} dx = \int_1^\infty x^{-\frac{3}{2}} dx = 2$$

اي ان التكامل متقارب وهذا يعني ان $E(\sqrt{X})$ موجود ومساوٍ الى 2 .

٢ ـ ١ ـ ٣ : خصائص التوقع الرياضي :

ان ماسبق توضيحه في الفقرتين السابقتين يقودنا للقول ان الرمز E يمثل عملية حساب متوسط اية دالة بدلالة متغير عشوائي X مثل $g(x)$ يسلك وفق دالة كتلة احتمالية او كثافة احتمالية . وفيما يلي بعض خصائص التوقع الرياضي وبفرض ان التوقع لاية دالة سترد في السياق معرف . مع ملاحظة ان هذه الخصائص قائمة سواء كان المتغير العشوائي متقطعاً أم مستمراً وعلى هذا الاساس فان البراهين ستورد لحالة المتغيرات المستمرة وهي ذاتها في حالة المتغيرات المتقطعة عدا انه يتم استبدال رمز التكامل برمز الجمع .

بفرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x)$ معرفة قيمة في Ω عندئذٍ

١ ـ اذا كانت k ثابتاً حقيقياً فان $E(K) = K$

البرهان :

$$E(K) = \int_\Omega K . f(x) dx = K \int f(x) dx$$

وحيث ان $f(x)$ دالة كثافة احتمالية فذلك يعني ان التكامل على Ω مساوٍ للواحد .

فاذن $E(K) = K(1) = K .$

٢ ـ اذا كان a ثابتاً حقيقياً وان $g(x)$ دالة بدلالة X فان $E[a.g(x)] = aE\,g(x)$

البرهان

$$E[a.g(x)] = \int_{\Omega} a\, g(x) f(x)\, dx$$

$$= a \int_{\Omega} g(x).f(x)\, dx = a\, E[g(x)]$$

٣ ـ اذا كان a,b ثابتين حقيقين وان $g(x)$ دالة بدلالة X فان

$$E[a\, g(x) + b] = a\, E[g(x)] + b$$

البرهان : يمكن برهنة ذلك وبسهولة باستخدام الخاصيتين (١ ، ٢) .

٤ ـ اذا كانت $a_1, a_2, \ldots, a_n$ ثوابت حقيقية وان $g_1(x), g_2(x), \ldots, g_n(x)$ دوال بدلالة X عندئذ

$$E\left[\sum_{i=1}^{n} a_i g_i(x)\right] = \sum_{i=1}^{n} a_i E g_i(x)$$

البرهان :

$$E\left[\sum_{i=1}^{n} a_i g_i(x)\right] = \int_{\Omega} \sum_{i=1}^{n} a_i g_i(x).f(x)\, dx$$

$$= \int_{\Omega} a_1 \cdot g_1(x) f(x)\, dx + \int_{\Omega} a_2 \cdot g_2(x) f(x)\, dx + \ldots + \int_{\Omega} a_n \cdot g_n(x) f(x)\, dx$$

$$= a_1 E g_1(x) + a_2 E g_2(x) + \ldots + a_n E g_n(x)$$

$$= \sum_{i=1}^{n} a_i E g_i(x)$$

٥ ـ بشكل عام اذا كانت $g(x)$ دالة بدلالة X فان $E[g(x)] \neq g[E(x)]$

وان $E\left(\frac{1}{g(x)}\right) \neq \frac{1}{E g(x)}$, $E\sqrt{g(x)} \neq \sqrt{E g(x)}$

كذلك فان

ثابت حقيقي $E \operatorname{Log} g(x) \neq \operatorname{Log} E g(x), E[g(x)]^k \neq [E g(x)]^k, k$

واذا كانت $g(x) = x$ فان

$$E\left(\frac{1}{X}\right) \neq \frac{1}{EX}, E\sqrt{X} \neq \sqrt{EX}, E \operatorname{Log} X \neq \operatorname{Log} EX$$

٦ ـ متباينة جينسون **Jenson's Inequality**

افرض ان $g(x)$ دالة مستمرة بدلالة المتغير العشوائي X عندئذ :

أ ـ اذا كانت $g(x)$ دالة محدبة(*) Convex function فان $E[g(x)] \geq g[EX]$

على سبيل المثال فان :

$EX^2 \geq [EX]^2$ طالما ان $g(x)=X^2$ دالة محدبة $(g''(x) = 2 > 0)$

$E\left(\frac{1}{X}\right) \geq \frac{1}{EX}$ ، $x > 0$ ، طالما ان $g(x) = \frac{1}{x}$ دالة محدبة

$\left(g''(x) = \frac{2}{x^3} > 0\right)$.

ب ـ اذا كانت $g(x)$ دالة مقعرة(**) Concave function فان $E[g(x)] \leq g[EX]$ على سبيل المثال فان :

$E\sqrt{X} \leq \sqrt{EX}$ ، $x>0$ ، طالما ان $g(x)=\sqrt{x}$ دالة مقعرة $(g''(x) < 0)$

$E(\operatorname{Log} X) \leq \operatorname{Log}(EX)$ ، $x > 0$ ، طالما ان $g(x) = \operatorname{Log} x$ دالة مقعرة .

(*) يقال ان الدالة المستمرة $g(x)$ محدبة في الفترة I اذا كان ، لاي عددين مثل a,b المعرفين في I ، $g\left(\frac{a+b}{2}\right) \leq \frac{1}{2} g(a) + \frac{1}{2} g(b)$ او بكلام آخر ، تكون الدالة محدبة اذا كانت $g''(x) > 0$

(**)
يقال ان الدالة $g(x)$ مقعرة في الفترة I اذا كان ، لاي عددين مثل a,b المعرفين في I ، $g\left(\frac{a+b}{2}\right) \geq \frac{1}{2} g(a) + \frac{1}{2} g(b)$ او بكلام آخر تكون الدالة $g(x)$ مقعرة اذا كانت $g''(x) < 0$

٢ ـ ١ ـ ٤ : تطبيقات التوقع الرياضي .

فيما يلي بعض تطبيقات التوقع الرياضي التي ستتم الحاجة لها في الكثير من الفقرات اللاحقة من فصول هذا الكتاب ، وهي :

اولاً : العزوم Moments

تعرف العزوم لمتغير عشوائي X (او لتوزيع احتمالي لمتغير عشوائي) بانها القيم المتوقعة لدوال معينة بدلالة X الذي يسلك وفق دالة كتلة احتمالية P(x) او دالة كثافة احتمالية f(x) . والعزوم على انواع عديدة منها مايلي :

ا ـ العزوم اللامركزية Non-central moments

افرض ان X متغير عشوائي وان a ثابت اختياري $a \neq EX$ عندئذٍ يقال ان التوزيع الاحتمالي الى X يمتلك عزماً لامركزياً ذا مرتبة r حول النقطة a معرف بالصيغة $E(X-a)^r, r = 1, 2, \dots$

ويتم حساب هذا العزم وفق الآتي :

في حالة X متقطع

$$E(X-a)^r = \sum_{x \in \Omega} (x-a)^r \cdot p(x)$$

في حالة X مستمر

$$= \int_{\Omega} (x-a)^r \cdot f(x)\, dx$$

مثال (٥) : افرض ان $p(x) = \frac{x}{10}, x = 1, 2, 3, 4$ تمثل دالة الكتلة الاحتمالية للمتغير X، جد العزم اللامركزي ذو المرتبة الثانية حول النقطة 2 .

الحل :

$$E(X-2)^2 = \sum_{x=1}^{4} (x-2)^2 \cdot \frac{x}{10}$$

$$= \frac{1}{10}\left[(-1)^2.(1) + (0)^2(2) + (1)^2(3) + (2)^2(4)\right]$$

$$= 2$$

مثال (٦) : لتكن $0 < x < 1, f(x)=2x$ تمثل دالة الكثافة الاحتمالية للمتغير X . جد العزم اللامركزي ذو المرتبة الثانية حول النقطة 2 .

الحل :

$$E(X-2)^2 = \int_0^1 (x-2)^2 . 2x\, dx$$

$$= 2\int_0^1 x(x-2)^2\, dx$$

$$= \frac{11}{6}$$

ب ـ العزوم حول نقطة الاصل Moments about the origin

ان هذا النوع من العزوم يعتبر حالة خاصة من العزوم اللامركزية في حالة اختيارنا $a = 0$ دائماً ووفق هذا الاختيار يقال ان التوزيع الاحتمالي الى X يمتلك عزم ذا مرتبة r حول نقطة الاصل معرف بالصيغة ... $EX^r, r = 1, 2,$ ويتم حساب هذا العزم وفق الآتي :

في حالة X متقطع $$EX^r = \sum_{x \in \Omega} x^r . P(x)$$

في حالة X مستمر $= \int_{\Omega} x^r . f(x)\, dx$

فاذا كانت $r = 1$ نحصل على العزم ذا المرتبة الاولى حول نقطة الاصل اي EX وهذا العزم يسمى الوسط Mean لقيم X في التوزيع الاحتمالي او القيمة المتوقعة الى X ، واذا كانت $r = 2$ نحصل على العزم ذا المرتبة الثانية حول نقطة الاصل اي EX^2 وهذا العزم يسمى الوسط لمربعات قيم X في التوزيع الاحتمالي .

مثال (٧) : لمعطيات المثال (٥) جد العزوم الثلاث الاولى حول نقطة الاصل .

الحل :

$$EX = \sum_{x=1}^{4} x . \frac{x}{10} = \frac{1}{10} (1^2 + 2^2 + 3^2 + 4^2) = 3$$

$$EX^2 = \sum_{x=1}^{4} x^2 . \frac{x}{10} = \frac{1}{10} (1^3 + 2^3 + 3^3 + 4^3) = 10$$

ويترك للقاريء حساب العزم الثالث .

مثال (٨) : لمعطيات المثال (٦) جد العزوم الثلاث الاولى حول نقطة الاصل .

الحل

$$EX = \int_0^1 x(2x)\, dx = 2 \int_0^1 x^2 dx = \frac{2}{3}$$

$$EX^2 = \int_0^1 x^2 (2x)\, dx = 2 \int_0^1 x^3 dx = \frac{1}{2}$$

ويترك للقاريء حساب العزم الثالث .

جـ ـ العزوم المركزية Central moments

وهذه هي الاخرى تعد حالة خاصة من العزوم اللامركزية في حالة اختيارنا $a = EX$ ووفق هذا الاختيار يقال ان التوزيع الاحتمالي الى X يمتلك عزماً مركزياً ذا مرتبة r معرفاً بالصيغة $E[X - EX]^r, r = 1, 2, \ldots$ ويتم حساب هذا العزم وفق الآتي :

$$E[X - EX]^r = \sum_{x \varepsilon \Omega} (x - EX)^r \cdot P(x) \qquad \text{في حالة X متقطع}$$

$$= \int_{\Omega} (x - EX)^r f(x)\, dx \qquad \text{في حالة X مستمر}$$

فاذا كانت $r = 1$ فان $E(X - E(X)) = 0$ (لماذا). واذا كانت $r = 2$ نحصل على العزم المركزي الثاني الذي يسمى التباين variance في التوزيع الاحتمالي الى X (او تباين X) الذي سيأتي ذكره في فقرة لاحقة.

مثال (٩) : لمعطيات المثال (٥) جد العزم المركزي الثالث .

الحل :

$$E(X - EX)^3 = EX^3 - 3(EX)(EX^2) + 2(EX)^3$$

واضح ان :

$$EX = 3, EX^2 = 10, EX^3 = 35{\cdot}4$$

فاذن

$$E(X - 3)^3 = -0{\cdot}6$$

مثال (١٠) : لمعطيات المثال (٦) جد العزم المركزي الثالث .

الحل :

واضح من معطيات هذا المثال ان :

$$EX = \frac{2}{3}, EX^2 = \frac{1}{2}, EX^3 = \frac{2}{5}$$

فاذن

$$E(X - EX)^3 = EX^3 - 3(EX)(EX^2) + 2(EX)^3$$

$$= -\frac{1}{135}$$

د ـ العزوم المطلقة المركزية Central Absolute moments

يقال ان التوزيع الاحتمالي الى X يمتلك عزماً مطلقاً مركزياً ذا مرتبة r معرف بالصيغة $E|X - EX|^r, r = 1, 2, \ldots$.ويتم حساب قيمة هذا العزم وفق الاتي :

$$E|X - EX|^r = \sum_{x \in \Omega} |x - EX|^r \cdot P(x)$$ في حالة X متقطع

$$= \int_{\Omega} |x - EX|^r \cdot f(x)\,dx$$ في حالة X مستمر

ويلاحظ لهذا النوع من العزوم انه اذا كانت r عدداً زوجياً فان العزوم المطلقة المركزية الزوجية المراتب ماهي الا العزوم المركزية الزوجية المراتب .
فاذا كانت $r = 2k, k = 1, 2, 3, \ldots$ فان

$$E|X - EX|^{2k} = E(X - EX)^{2k}$$

مثال (١١) : لمعطيات المثال (٥) جد العزم المطلق المركزي الثالث .
الحل :

$$E|X - EX|^3 = E|X - 3|^3 = \sum_{x=1}^{4} |X - 3|^3 \cdot \frac{x}{10}$$

$$= \frac{1}{10}[(2)^3(1) + (1)^3(2) + (0)^3(3) + (1)^3(4)]$$

$$= 0.8$$

مثال (١٢) : لمعطيات المثال (٦) جد العزم المطلق المركزي الثالث .

الحل :

$$E|X - EX|^3 = E|X - \frac{2}{3}|^3 = \int_0^1 |x - \frac{2}{3}|^3 \cdot 2x\,dx$$

$$= 2\int_0^{2/3} - x\left(x - \frac{2}{3}\right)^3 dx + 2\int_{2/3}^1 x\left(x - \frac{2}{3}\right)^3 dx = \frac{133}{1215}$$

هـ ـ العزوم العاملية Factorial moments

يقال ان التوزيع الاحتمالي الى X يمتلك عزماً عاملياً ذا مرتبة r معرفاً بالصيغة $r = 1, 2, \ldots$ ، $E\left[\prod_{j=1}^{r}(X - j + 1)\right]$ ويتم حساب هذا العزم وفق الآتي:

في حالة X متقطع

$$E\left[\prod_{j=1}^{r}(X - j + 1)\right] = \sum_{x\in\Omega}\left[\prod_{j=1}^{r}(x - j + 1)\right]P(x)$$

في حالة X مستمر

$$= \int_{\Omega}\left[\prod_{j=1}^{r}(x - j + 1)\right]f(x)\,dx$$

واذا كانت $r = 1$ فان العزم العاملي ذو المرتبة الاولى ماهو الا EX ، وعندما $r = 2$ فان العزم العاملي ذو المرتبة الثانية ماهو الا $EX(X - 1) = EX^2 - EX$

مثال (١٣) : لمعطيات المثال (٦) جد العزم العاملي الثالث .

الحل :

$$EX(X - 1)(X - 2) = EX^3 - 3EX^2 + 2EX$$

ومن معطيات هذا المثال لاحظنا أن :

$$EX^3 = \frac{2}{5}, \quad EX^2 = \frac{1}{2}, \quad EX = \frac{2}{3}$$

عليه فان

$$EX(X-1)(X-2) = \frac{7}{30}$$

و ــ العلاقة بين العزوم المركزية والعزوم حول نقطة الاصل .

فيما يلي بعض العلاقات التي تربط ما بين العزوم المركزية والعزوم حول نقطة الاصل . هذه العلاقات مفيدة من الناحية التطبيقية عند حساب عزوم مركزية لتوزيع معين علمت فيه مسبقاً عزوم حول نقطة الاصل . ان العزم المركزي ذو المرتبة r هو $E(X-EX)^r$. وباستخدام نظرية ثنائي الحدين Binomial theorem يمكن البيان ان

$$(X-EX)^r = \sum_{k=0}^{r} C_k^r X^k (-EX)^{r-k}$$

$$= (-EX)^r + rX(-EX)^{r-1} + \dots + X^r$$

$$\therefore \; E(X-EX)^r = \sum_{k=0}^{r} C_k^r (-EX)^{r-k} . EX^k , \; EX^0 = 1$$

لاحظ من الصيغة الاخيرة أنه امكن التعبير عن العزم المركزي ذا المرتبة r بدلالة العزوم حول نقطة الاصل . فمثلاً

$$E(X-EX)^3 = \sum_{k=0}^{3} C_k^3 (-EX)^{3-k} . EX^k$$

$$= (-EX)^3 + 3(-EX)^2(EX) + 3(-EX)(EX^2) + EX^3$$
$$= EX^3 - 3(EX)(EX^2) + 3(EX)^3 - (EX)^3$$
$$= EX^3 - 3(EX)(EX^2) + 2(EX)^3$$

ز ـ العلاقة بين العزوم العاملية والعزوم حول نقطة الاصل

يمكن حساب قيمة العزوم العاملية لتوزيع احتمالي اذا علمت عزومه حول نقطة الاصل . فمثلاً يمكن حساب قيمة العزم العاملي الثالث بدلالة العزوم الثلاثة الاولى حول نقطة الاصل . اي ان

$$E\left[\prod_{j=1}^{3}(X-j+1)\right] = EX(X-1)(X-2)$$

$$= EX^3 - 3EX^2 + 2EX$$

وخير مثال للفائدة التطبيقية للعلاقات ما بين العزوم هو الموضح بالمثالين المرقمين (٩) . (١٣) .

ثانياً : المتوسط Mean

ويسمى في بعض الاحيان « الوسط الحسابي لقيم المتغير العشوائي في التوزيع الاحتمالي » . ويعرف المتوسط بأنه قيمة العزم ذو المرتبة الاولى حول نقطة الاصل . وغالباً ما يرمز لهذا المؤشر بالرمز μ_x او بشكل مختصر μ. وهذا يعني ان :

$$\mu = EX$$

$$= \sum_{x\varepsilon\Omega} x.\,P(x) \qquad \text{في حالة X متقطع}$$

$$= \int_{\Omega} xf(x)\,dx \qquad \text{في حالة X مستمر}$$

ان المتوسط عبارة عن مقياس موقعي (اي قيمةمعرفة على المحور السيني) مقاس بنفس وحدات قياس المتغير X ويعبر عن قيم X المعرفة في Ω بقيمة واحدة تفي ببعض المعلومات عن موقع التوزيع الاحتمالي . وفي حالة المتغيرات المستمرة فان $\mu \in \Omega$ اما في حالة المتغيرات المتقطعة فان قيمة μ قد تكون معرفة في Ω

او قد لاتكون الا انه وبشكل عام اذا كان Ω مجموعة قابلة للعد فان μ يجب ان يكون اكبر من اصغر قيمة معرفة في Ω واصغر من اكبر قيمة معرفة في Ω . فمثلاً اذا كان $\Omega = \{ x : x = 0, 1, 2, \ldots \}$ فان $\mu \in \{ 0 < x < \infty \}$. وتجدر الاشارة هنا الى انه في بعض الحالات لايمكن ايجاد المتوسط لتوزيع احتمالي معين بسبب عدم تحقق خاصية التقارب المطلق اي في حالة كون :

$$\sum_{x \in \Omega} |x| P(x) \to \infty \qquad \int_{\Omega} |x| f(x)\, dx \to \infty$$

وبشكل عام اذا كان $x \in \Omega = \{ x : -\infty < x < \infty \}$ فان $-\infty < \mu < \infty$ اذا كان μ موجود .

ان اهم خصائص هذا المؤشر هي :

أ ـ ان $E(X - \mu) = 0$

البرهان : $E(X - \mu) = EX - E\mu = \mu - \mu = 0$

ب ـ ان $E(X - b)^2$, b عدد حقيقي ، يكون اقل مايمكن عندما $b = \mu$

البرهان

$$\begin{aligned} E(X - b)^2 &= E(X - b + \mu - \mu)^2 \\ &= E[(X - \mu) - (b - \mu)]^2 \\ &= E(X - \mu)^2 + E(b - \mu)^2 - 2E(X - \mu)(b - \mu) \\ &= E(X - \mu)^2 + (b - \mu)^2, (b - \mu)E(X - \mu) = 0 \end{aligned}$$

واضح ان $(b - \mu)^2 \geq 0$ ويكون مساو للصفر عندما $b = \mu$ وذلك يعني ان $E(X - b)^2$ يكون اقل مايمكن عندما $b = \mu$

مثال (١٤) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{1}{8}$, $x = 1, 2, \ldots, 8$ جد μ .

الحل :

$$\mu = \sum_{x=1}^{8} x.\ \frac{1}{8} = \frac{1}{8} \sum_{x=1}^{8} x = \frac{36}{8} = 4.5$$

لاحظ في هذا المثال ان μ غير معرف في Ω الا ان $1 < \mu < 8$

مثال (١٥) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{4^{x} e^{-4}}{x!}$, $x = 0, 1, 2, \ldots$ جد متوسط X في هذا التوزيع .

الحل :

$$\mu = \sum_{0=x}^{\infty} x.\ \frac{4^{x} e^{-4}}{x!} = 4e^{-4} \sum_{x=1}^{\infty} \frac{4^{x-1}}{(x-1)!}$$

$$= 4e^{-4}.\ e^{4} = 4$$

لاحظ في هذا المثال ان $\mu = 4$ قيمة معرفة في Ω

مثال (١٦) : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = 3e^{-3x}$, $x \geq 0$ جد متوسط X في هذا التوزيع .

الحل :

$$\mu = \int_{0}^{\infty} x.\ 3e^{-3x}dx = 3 \int_{0}^{\infty} xe^{-3x}dx$$

وبحل التكامل بطريقة التجزئة نلاحظ ان

$$\int_{0}^{\infty} xe^{-3x}dx = \frac{1}{9} \quad \therefore \mu = \frac{1}{3}$$

مثال (١٧) : افرض ان $f(x) = \frac{1}{x^2}, 1 < x < \infty$ تمثل دالة الكثافة الاحتمالية الى X جد μ_x .

الحل :

$$\mu_x = \int_1^{\infty} x . \frac{1}{x^2} dx = \int_1^{\infty} \frac{1}{x} dx = \ln x \Big]_1^{\infty}$$

$$= \lim_{x \to \infty} (\ln x)$$

لاحظ هنا ان التكامل متباعد وهذا يعني انه لايمكن ايجاد المتوسط لهذا التوزيع .

ثالثاً: الوسط التوافقي Harmonic mean

يعرف الوسط التوافقي لقيم X في توزيع احتمالي معين بانه مقلوب القيمة المتوقعة لمقلوب X. فاذا كان H يمثل الوسط التوافقي فان

$$H = \frac{1}{E\left(\frac{1}{X}\right)} \quad \text{أو} \quad \frac{1}{H} = E\left(\frac{1}{X}\right)$$

وهذا يعني ان

$$\frac{1}{H} = \sum_{x \in \Omega} \frac{1}{x} P(x), x \neq 0 \qquad \text{في حالة X متقطع}$$

$$= \int_{\Omega} \frac{1}{x} f(x) dx \qquad \text{في حالة X مستمر}$$

وبشكل عام اذا كان $x \in \Omega = \{x : -\infty < x < \infty\}$ فان $-\infty < H < \infty$ كذلك ومن خلال متباينة جينسون يمكن ملاحظة انه اذا كان $x > 0$ فان

$$\frac{1}{H} = E\left(\frac{1}{X}\right) \geq \frac{1}{EX}, EX \neq 0$$

وهذا يعني ان $H \leq \mu_x$ ان مسألة ايجاد الوسط التوافقي لتوزيع احتمالي مرهونة بكون ان $E\left(\frac{1}{X}\right)$ يتمتع بخاصية التقارب المطلق .

مثال (١٨) : افرض ان $P(x) = \frac{1}{8}\ , x = 1, 2, \ldots, 8$ تمثل دالة الكتلة الاحتمالية الى X جد الوسط التوافقي لهذا التوزيع .

الحل :

$$\frac{1}{H} = E\left(\frac{1}{X}\right) = \sum_{x=1}^{8} \frac{1}{x} \cdot \frac{1}{8} = \frac{1}{8} \sum_{x=1}^{8} \frac{1}{x} = \frac{2{\cdot}7179}{8}$$

$$= 0{\cdot}3397 \qquad \therefore H = 2{\cdot}9438 < \mu_x = 4{\cdot}5$$

مثال (١٩) : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = 2x$ $0 < x < 1$, جد الوسط التوافقي لهذا التوزيع .

الحل :

$$\frac{1}{H} = \int_0^1 \frac{1}{x} \cdot 2x\,dx = 2$$

فاذن
$$H = \frac{1}{2} < \mu_x = \frac{2}{3}$$

مثال (٢٠) : ليكن X متغير عشوائي بدالة كثافة احتمالية $f(x) = \frac{1}{x^2}$ $1 < x < \infty$, جد الوسط التوافقي لهذا التوزيع .

الحل :

$$\frac{1}{H} = \int_{1}^{\infty} \frac{1}{x} \cdot \frac{1}{x^2}\, dx = \int_{1}^{\infty} \frac{1}{x^3}\, dx = \frac{1}{2}$$

فاذن $H = 2$

لاحظ من هذا المثال ان على الرغم من ان μ_x غير موجود الا ان التوزيع امتلك وسطاً توافقياً .

رابعاً : التباين Variance

ان التباين عبارة عن مقياس لدرجة تشتت قيم المتغير العشوائي لتوزيع احتمالي معين . ويعرف التباين بانه قيمة العزم المركزي الثاني . وغالباً مايرمز لتباين قيم المتغير X بالرمز $V(X)$ او σ_x^2 وهذا يعني ان

$$\sigma_x^2 = E[X - EX]^2 = EX^2 - (EX)^2$$

اي ان التباين ماهو الا الفرق مابين العزم الثاني حول نقطة الاصل ومربع العزم الاول حول نقطة الاصل . اي انه « متوسط » مربع الفرق ما بين قيم X ومتوسطها . وحسب متباينة جينسون فان $EX^2 \geq (EX)^2$ وذلك يعني ان σ_x^2 قيمة غير سالبة أي أن $\sigma_x^2 \geq 0$. أي أنه وبشكل عام اذا كان $x \in \Omega = \{x : -\infty < x < \infty\}$ فان $\sigma_x^2 \geq 0$. ان الجذر التربيعي الموجب للتباين، أي σ_x ، يسمى «الانحراف المعياري Standard deviation» لقيم X. ان التباين مقاس بمربع وحدات قياس المتغير X (سم٢، م٢، غم٢) في حين ان الانحراف المعياري يكون مقاساً بنفس وحدات X . ان مسألة تحديد التباين الى X في توزيع احتمالي معين مرهون بتحديد قيمة العزمين الاول والثاني حول نقطة الاصل وفي حالة عدم تمتع احدهما بخاصية التقارب المطلق فذلك يعني عدم امكانية ايجاد قيمة التباين . وفيما يلي بعض خصائص هذا المؤشر :

أ ــ اذا كانت a كمية ثابتة فان $V(a) = 0$

البرهان :

$$V(a) = Ea^2 - (Ea)^2 = a^2 - (a)^2 = a^2 - a^2 = 0$$

ب ـ اذا كانت a كمية ثابتة . X متغيراً عشوائياً . وان $Y = aX$ عندئذٍ
$$V(Y) = a^2V(X)$$

البرهان :

$$V(Y) = EY^2 - (EY)^2 = Ea^2X^2 - E(aX)^2$$
$$= a^2(EX^2 - (EX)^2) = a^2V(X)$$

جـ ـ اذا كانت كل من a . b كمية ثابتة . X متغير عشوائي وان $Y = aX \pm b$
عندئذٍ $V(Y) = a^2V(X)$

البرهان : باستخدام الخاصيتين (ا . ب) يمكن برهنة ذلك .

د ـ ليكن σ_x^2, μ_x يمثلان على التوالي الوسط والتباين لقيم X في توزيع احتمالي معين .وافرض ان $Z = (X - \mu_x) / \sigma_x$ حيث Z تسمى « الدرجة المعيارية Standard score » المقابلة الى X . عندئذٍ $EZ = 0$
$V(Z) = 1$ مهما كان التوزيع الاحتمالي الى X .

البرهان :

$$EZ = E(X - \mu x) / \sigma_x = \frac{1}{\sigma_x} E(X - \mu_x) = 0$$

$$V(Z) = EZ^2 - (EZ)^2 = EZ^2 = E\left(\frac{X - \mu_x}{\sigma_x}\right)^2$$

$$= \frac{1}{\sigma_x^2} E(X - \mu_x)^2 = \frac{1}{\sigma_x^2} \cdot \sigma_x^2 = 1$$

مثال (٢١) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{x}{6}$,
$x = 1,2,3$. جد تباين X وتباين $Y = 4X$

الحل :

$$EX = \sum_{x=1}^{3} x \cdot \frac{x}{6} = \frac{1}{6} \sum_{x=1}^{3} x^2 = \frac{7}{3}$$

$$EX^2 = \sum_{x=1}^{3} x^2 . \frac{x}{6} = \frac{1}{6} \sum_{x=1}^{3} x^3 = 6$$

$$\therefore \; V(X) = 6 - \left(\frac{7}{3} \right)^2 = \frac{5}{9}$$

$$V(Y) = V(4X) = 16V(X) = 16 . \frac{5}{9} = \frac{80}{9}$$

مثال (٢٢) : لتكن $f(x) = 2e^{-2x}, x \geq 0$ تمثل دالة الكثافة الاحتمالية الى X
جد تباين X وتباين $Y = 3 - 2X$

الحل :

$$EX = \int_0^x 2xe^{-2x}dx = \frac{1}{2} , EX^2 = \int_0^x 2x^2e^{-2x}dx = \frac{1}{2}$$

فاذن
$$V(X) = \frac{1}{2} - \left(\frac{1}{2} \right)^2 = \frac{1}{4}$$

وان
$$V(Y) = V(3 - 2X) = V(-2X)$$

$$= 4V(X) = 4 \cdot \frac{1}{4} = 1$$

خامساً : الانحراف المطلق Mean deviation (M_d)

مقياس لدرجة تشتت قيم المتغير العشوائي في توزيع احتمالي . ويعرف الانحراف المطلق بانه قيمة العزم المركزي المطلق الاول . وهذا يعني ان

$$M_d = E|X - EX|$$

اي ان الانحراف المطلق يمثل متوسط الفرق المطلق بين EX , X . ووفق هذا التعريف نلاحظ ان M_d قيمة غير سالبة . اي $M_d \geq 0$. وبشكل عام اذا كان $x \in \Omega = \{x: -\infty < x < \infty\}$ فان $M_d \geq 0$. كذلك يلاحظ ان M_d مقاس بنفس وحدات قياس المتغير X . ان مسألة ايجاد M_d لقيم في توزيع احتمالي مرهونة بايجاد EX . وهذا يعني أنه يجب ان يكون EX متقارب على نحو مطلق كي يسمح لنا ذلك حساب قيمة M_d . كذلك فان من خواص هذا المقياس هي: اذا كانت كل من a ، b كمية ثابتة $Y = aX \pm b$ فان الانحراف المطلق الى Y هو $M_d(Y) = |a| . M_d(X)$.

مثال (٢٣) : افرض ان $P(x) = \frac{1}{7}, x = 1, 2, ..., 7$ تمثل دالة الكتلة الاحتمالية إلى X . جد الانحراف المطلق .

الحل : حيث ان $EX = 4$ فان $M_d = E|X - EX| = E|X - 4|$

$$= \sum_{x=1}^{7} |X - 4| . \frac{1}{7} = \frac{1}{7}(3 + 2 + 1 + 0 + 1 + 2 + 3) = \frac{12}{7}$$

مثال (٢٤) : لتكن $f(x) = \frac{1}{10}, -3 < x < 7$ تمثل دالة الكثافة الاحتمالية الى X . جد الانحراف المطلق الى X . الانحراف المطلق الى $Y = 4 - 3X$

الحل : نجد اولا EX وهي

$$EX = \int_{-3}^{7} x . \frac{1}{10} dx = \frac{1}{10} \int_{-3}^{7} x dx = 2$$

فاذن

$$M_d(X) = E|X-2| = \int_{-3}^{7} |x-2| \cdot \frac{1}{10}\, dx$$

$$= \frac{1}{10}\left[\int_{-3}^{2} -(x-2)\,dx + \int_{2}^{7} (x-2)\,dx\right] = 2{\cdot}5$$

وان

$$M_d(Y) = M_d(4-3X) = |-3| \cdot M_d(X) = 3(2{\cdot}5) = 7{\cdot}5$$

٢ ـ ٢ : الدوال المولدة للعزوم

Moment generating functions

استعرضنا في الفقرة السابقة مفهوم التوقع الرياضي وخصائصه واهم تطبيقاته وخصوصاً في مجال حساب عزوم المتغير العشوائي في توزيع احتمالي . في هذه الفقرة سوف نستعرض دوال معينة من شأنها توليد عزوم توزيع احتمالي بانواعها المختلفة اياً كانت مراتبها . مع ملاحظة ان البراهين التي سترد لاحقاً في هذه الفقرة ستكون لحالة المتغيرات العشوائية المستمرة وهي ذاتها لحالة المتغيرات العشوائية المتقطعة ماعدا انه يتم استبدال رمز التكامل برمز الجمع .

افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x)$ وان Ω بشكل عام تمثل مجموعة الاعداد الحقيقية R . وافرض ان $g(x)$ دالة بدلالة X وان t متغير آخر و h عدد موجب بحيث ان $-h < t < h$.وفق هذه المعطيات يمكن تعريف الدالة المولدة لعزوم $g(x)$ (اذا كانت موجودة) في هذا التوزيع بانها القيمة المتوقعة الى الدالة $e^{t \cdot g(x)}$ فاذا رمزنا للدالة المولدة لعزوم $g(x)$ بالرمز $M_{g(x)}(t)$ فذلك يعني ان :

$$M_{g(x)}(t) = Ee^{t \cdot g(x)} = \sum_{x \epsilon \Omega} e^{t \cdot g(x)} \cdot p(x)$$ في حالة X فتقطع

$$= \int_{-\infty}^{+\infty} e^{t \cdot g(x)} \cdot f(x)\, dx$$ في حالة X مستمر

ان مسألة وجود الدالة المولدة لعزوم g (x) مرهون بكون التكامل او المجموع متقارب على نحو مطلق واذا لم يكن كذلك عندئذٍ يقال ان الدالة المولدة لعزوم g (x) غير موجودة . واذا كانت هذه الدالة موجودة عندئذٍ يمكن التعرف على عزوم التوزيع الاحتمالي الذي اشتقت منه مهما كان نوع تلك العزوم او مراتبها وكما هو موضح بالآتي :

باستخدام مفكوك سلسلة مكلورين Maclaurin's expansion يمكن ملاحظة ان :

$$e^{t\cdot g(x)} = \sum_{k=0}^{\infty} \frac{[t\cdot g(x)]^k}{k!}$$

$$= 1 + t\cdot g(x) + \frac{[t\cdot g(x)]^2}{2!} + \frac{[t\cdot g(x)]^3}{3!} + \ldots$$

وبأخذ التوقع الرياضي لطرفي السلسلة نحصل على :

$$M_{g(x)}(t) = 1 + t\,Eg(x) + \frac{t^2}{2!}Eg^2(x) + \frac{t^3}{3!}Eg^3(x) + \ldots * $$

يلاحظ من (*) ان عزوم الدالة g (x) موجودة ولمختلف المراتب . وهذا يعني انه يمكن توليد اي عزم منها فيما لو تم ايجاد المشتقة من مرتبة ذلك العزم للدالة M (t) نسبة الى t ومن ثم جعل t مساوية للصفر وكما هو مبين بالآتي :

$$M'_{g(x)}(t) = E\,g(x) + t\,Eg^2(x) + \frac{t^2}{2!}Eg^3(x) + O'(t) \ldots (*')$$

حيث o'(t) تشير الى مشتقات من المرتبة الاولى لحدود لاحقة تتضمن t بقوى عليا . وبجعل t = 0 في (*') نحصل على :

$$M'_{g(x)}(o) = E\,g(x)$$

كذلك فان

$$M''_{g(x)}(t) = Eg^2(x) + t\,Eg^3(x) + o''(t) \qquad \ldots (*'')$$

حيث $(t)''0$ تشير الى مشتقات من المرتبة الثانية لحدود لاحقة تتضمن t بقوى عليا . وبجعل o = t في (*") نحصل على :

$$M''_{g(x)}(o) = g^2(x)$$

ووفق هذا الاجراء فان العزم ذو المرتبة r للدالة g (x) ماهي الا المشتقة ذات المرتبة r للدالة M وبجعل t مساوية للصفر . اي ان

$$M^{(r)}_{g(x)}(o) = Eg^r(x), r = 1, 2, 3, \ldots$$

وفيما يلي الانواع المختلفة للدوال المولدة للعزوم وسوف نركز اهتمامنا على الدالة المولدة للعزوم حول نقطة الاصل نظراً لاهميتها .

٢ ـ ٢ ـ ١ : الدالة المولدة للعزوم (حول نقطة الاصل) Moment generating function (about the origion)

في حالة اختيار الدالة $g(x) = x$ في (*) نحصل على ماتسمى بالدالة المولدة لعزوم X حول نقطة الاصل . ووفقا لهذا الاختيار فان هذه الدالة (اذا كانت موجودة) سوف تكون معرفة بالآتي :

في حالة X متقطع
$$M_X(t) = Ee^{tX} = \sum_{x\varepsilon\Omega} e^{tx} \cdot p(x)$$

في حالة X مستمر
$$= \int_{-\infty}^{+\infty} e^{tx} \cdot f(x)\,dx$$

واضح مما تقدم ان $M_X(0) = 1$. وبذلك فان العزم ذو المرتبة r حول نقطة الاصل ماهو الا :

$$EX^r = M^{(r)}_X(0), r = 1, 2, 3, \ldots$$

وهذا يعني ان :

متوسط قيم X في التوزيع هو $M'_X(0) = EX$,

وان

متوسط مربعات قيم X في التوزيع هو : $M''_X(0) = EX^2$

وهذا يعني ان :

$$V(X) = M''_X(0) - [M'_X(0)]^2$$

وفيما يلي بعض خصائص هذه الدالة :

١ ـ بفرض ان X متغير عشوائي وان $M_X(t)$ موجودة . وبفرض ان $Y = a + bX$ حيث a و $b \neq 0$ ثابتان حقيقيان عندئذٍ

$$M_Y(t) = e^{at} . M_X(bt)$$

البرهان : نفرض ان $M_Y(t)$ تمثل الدالة المولدة لعزوم Y حول نقطة الاصل .

عندئذٍ

$$\begin{aligned} M_Y(t) &= Ee^{tY} = Ee^{t(a+bX)} \\ &= Ee^{at} . e^{(bt)X} = e^{at} . Ee^{(bt)X} \\ &= e^{at} M_X(bt) \end{aligned}$$

٢ ـ ان الدالة المولدة لعزوم الدرجة المعيارية Z حول نقطة الاصل هي :

$$M_Z(t) = e^{-\frac{\mu}{\sigma} t} M_X\left(\frac{t}{\sigma}\right)$$

البرهان : نفرض ان $M_Z(t)$ موجودة وان $Z = (X - \mu)/\sigma$ تمثل الدرجة المعيارية المقابلة الى X . عندئذٍ فان

$$M_Z(t) = Ee^{tZ} = Ee^{t\left(\frac{X-\mu}{\sigma}\right)}$$

$$= Ee^{\frac{t}{\sigma} X} . e^{-\frac{\mu}{\sigma} t} = e^{-\frac{\mu}{\sigma} t} Ee^{\frac{t}{\sigma} X}$$

$$= e^{-\frac{\mu}{\sigma} t} . M_X\left(\frac{t}{\sigma}\right)$$

٣ ـ اذا كانت $M_X(t)$ موجودة فان $K_X(t) = \ln M_X(t)$ موجودة . حيث $K_X(t)$ تسمى « الدالة المولدة التراكمية " Cumulant generating function فان $K''_X(0) = V(X),\ K'_X(0) = EX$

البرهان :

$$K'_X(t) = \frac{M'_X(t)}{M_X(t)} \qquad \therefore K'_X(0) = \frac{M'_X(0)}{M_X(0)}$$

وحيث ان $M_X(0) = 1$ فذلك يعني أن :

$$K'_X(0) = M'_X(0) = EX$$

كذلك فان

$$K''_X(t) = \frac{M_X(t).M''_X(t) - M'_X(t).M'_X(t)}{M^2_X(t)}$$

فاذن

$$K''_X(0) = M''_X(0) - [M'_X(0)]^2 = V(X)$$

مع ملاحظة ان

$$K_X(0) = \ln M_X(0) = 0$$

٤ ـ اذا كانت $M_X(t)$ موجودة وان EX معرف . عندئذٍ وحسب متباينة جينسون فان $M_X(t)$ تكون في نهايتها الصغرى عندما $t = 0$ وعند ذلك فان هذه الدالة سوف تمتلك حد ادنى مقداره واحد .

البرهان : ان $M_X(t) = Ee^{tX}$ وحسب متباينة جينسون فان $Ee^{tX} \geq e^{tEX}$ طالما ان e^{tX} دالة محدبة بالمتغير X لجميع قيم t وقيم x واذا كان $EX = 0$ فان $Ee^{tX} \geq 1$. كما ان هذا الحد يمكن الوصول اليه عندما $t = 0$. وهذا يعني ان $M_X(t)$ تكون في نهايتها الصغرى عندما $t = 0$ وتمتلك حد ادنى مقداره واحد .

٥ ـ اذا كانت $M_X(t)$ موجودة فهي دالة وحيدة unique function لذلك التوزيع الاحتمالي . اي ان كل توزيع احتمالي امتلك دالة مولدة للعزوم فهي دالة وحيدة لايوجد غيرها . وهذا يعني ان $M_X(t)$ تصف التوزيع الاحتمالي الى X (اي من هو التوزيع الاحتمالي الذي اشتقت منه $M_X(t)$) والعكس صحيح اي ان التوزيع الاحتمالي يصف الدالة المولدة للعزوم اذا كانت موجودة . وسوف نلاحظ في الفقرة (٧ ـ ٣) اسلوب استنتاج التوزيعات الاحتمالية باستخدام هذه الدالة .

مثال (٢٥): افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{3^x e^{-3}}{x!}$,
$x = 0, 1, 2, \ldots$. جد مايلي :

أ ـ الدالة المولدة لعزوم X حول نقطة الاصل .
ب ـ الدالة المولدة التراكمية .
جـ ـ الوسط والتباين باستخدام الدالتين المستخرجتين في أ . ب .
د ـ الدالة المولدة لعزوم $Y = 2 - 3X$.

الحل : نفرض ان $M_X(t)$ موجودة . عندئذٍ :

أ ـ
$$M_X(t) = Ee^{tX} = \sum_{x=0}^{x} e^{tx} . \frac{3^x e^{-3}}{x!}$$

$$= e^{-3} \sum_{x=0}^{x} \frac{(3e^t)^x}{x!} = e^{-3} . e^{3e^t} \neq e^{3(e^t - 1)}$$

لاحظ ان $M_X(0) = 1$.

ب ـ
$$K_X(t) = \ln M_X(t) = 3(e^t - 1)$$

لاحظ ان $K_X(0) = 0$.

جـ ـ الوسط والتباين باستخدام $M_X(t)$

$$M'_X(t) = e^{3(e^t - 1)} . 3e^t = 3e^{3(e^t - 1) + t}$$

$$\therefore EX = M'_X(0) = 3$$
$$M''_X(t) = 3(e^{3(e^t - 1) + t}(3e^t + 1))$$
$$\therefore EX^2 = M''_X(0) = 12 \qquad \therefore V(X) = 12 - 3^2 = 3$$

الوسط والتباين باستخدام $K_X(t)$

$$K'_X(t) = 3e^t \qquad \therefore EX = K'_X(0) = 3$$
$$K''_X(t) = 3e^t \qquad \therefore V(X) = K''_X(0) = 3$$

لاحظ السهولة في حساب الوسط والتباين باستخدام الدالة المولدة التراكمية .

د ـ

$$M_Y(t) = e^{at} . M_X(bt) = e^{2t} . e^{3(e^{-3t}-1)}$$

مثال (٢٦) : افرض ان $\grave{X}$ متغير عشوائي بدالة كثافة احتمالية $f(x) = 2e^{-2x}, x \geq 0$ ، جد مايلي :

أ ـ الدالة المولدة للعزوم حول نقطة الاصل .

ب ـ الدالة المولدة التراكمية .

جـ ـ الوسط والتباين الى X .

د ـ الدالة المولدة لعزوم الدرجة المعيارية في هذا التوزيع .

الحل : نفرض ان $M_X(t)$ موجودة . فاذن

أ ـ

$$M_X(t) = Ee^{tX} = \int_0^\infty e^{tx} . 2e^{-2x} dx$$

$$= 2\int_0^\infty e^{-x(2-t)} dx = \frac{2}{t-2}\left[e^{-x(2-t)}\right]_0^\infty = \frac{2}{2-t}, t < 2$$

ب ـ

$$K_X(t) = \ln M_X(t) = \ln 2 - \ln(2-t)$$

جـ ـ

$$K_X'(t) = \frac{1}{2-t} \qquad \therefore EX = K_X'(0) = \frac{1}{2}$$

$$K_X''(t) = \frac{1}{(2-t)^2} \quad \therefore V(X) = K_X''(0) = \frac{1}{4} \quad \therefore \sigma_X = \frac{1}{2}$$

د ـ

$$M_z(t) = e^{-\frac{\mu}{\sigma}t} M_X\left(\frac{t}{\sigma}\right)$$

$$= e^{-t} . \frac{2}{2-2t} = \frac{e^{-t}}{1-t}, t < 1$$

مثال (٢٧) : افرض ان $f(x) = \frac{1}{x^2}, x \geq 1$ تمثل دالة الكثافة الاحتمالية الى X . جد الدالة المولدة لعزوم X حول نقطة الاصل .

الحل : نفرض ان $M_X(t)$ موجودة . فاذن :

$$M_X(t) = Ee^{tX} = \int_1^{\infty} e^{tx} \cdot \frac{1}{x^2} dx$$

وباستخدام اسلوب التكامل بالتجزئة وبمرحلتين يمكن ملاحظة ان :

$$\int_1^{\infty} x^{-2} \cdot e^{tx} dx = \left[\lim_{x \to \infty} (-x^{-1} \cdot e^{tx}) + e^t \right] + \left[\lim_{x \to \infty} te^{tx} \ln x \right]$$

$$- t^2 \int_1^{\infty} e^{tx} \ln x dx$$

واضح ان التكامل اعلاه متباعد وهذا يعني ان $M_X(t)$ غير موجودة . فاذن نستنتج ان هذا التوزيع الاحتمالي لايمتلك دالة مولدة للعزوم حول نقطة الاصل الامر الذي يقودنا للقول انه لايمكن ايجاد الوسط والتباين الى X .

٣ ـ ٢ ـ ٢ : الدالة المولدة للعزوم اللامركزية
Non-central moment generating function

اذا تم اختيار الدالة $g(x) = x - a$ حيث a ثابت اختياري . عندئذٍ يمكن توليد العزم اللامركزي ذا المرتبة r من خلال التعويض عن $g(x)$ بـ $(x - a)$ في الصيغة (*) . اي ان الدالة المولدة للعزوم اللامركزية ستكون :

$$M_{(X-a)}(t) = Ee^{t(X-a)} = e^{-at} \cdot M_X(t)$$

وبذلك فان العزم اللامركزي ذو المرتبة r حول النقطة a هو :

$$E(X - a)^r = M^{(r)}_{(X-a)}(t) \quad , r = 1, 2, 3, \ldots$$

يلاحظ فيما تقدم ان وجود الدالة المولدة للعزوم اللامركزية مرتبط بوجود الدالة المولدة للعزوم حول نقطة الاصل .

مثال (٢٨) : اذا علمت ان الدالة المولدة لعزوم X حول نقطة الاصل هي

$$M_X(t) = e^{\frac{1}{2}t^2}$$

جد الدالة المولدة للعزوم اللامركزية حول النقطة 3 ثم جد العزم اللامركزي الثاني .

الحل : حيث ان $M_X(t)$ موجودة فذلك يعني ان الداله المولدة للعزوم اللامركزية موجودة ايضاً وهي :

$$M_{(X-3)}(t) = e^{-3t}.e^{\frac{1}{2}t^2} = e^{-3t+\frac{1}{2}t^2}$$

فاذن :

$$M'_{(X-3)}(t) = e^{-3t+\frac{1}{2}t^2}.(-3+t)$$

$$\therefore M'_{(X-3)}(0) = E(X-3) = -3$$

$$M''_{(X-3)}(t) = e^{-3t+\frac{1}{2}t^2}.(1) + (-3+t)^2.e^{-3t+\frac{1}{2}t^2}$$

$$M''_{(X-3)}(0) = E(X-3)^2 = 1 + (-3)^2 = 10$$

٣ ـ ٢ ـ ٣ : الدالة المولدة للعزوم المركزية

central moment generating function

ان هذه الدالة تعد حالة خاصة من الدالة المولدة للعزوم اللامركزية عند اختيار $a = \mu_x = \mu$. وعندئذٍ فان هذه الدالة تأخذ الشكل التالي :

$$M_{(X-\mu)}(t) = e^{-\mu t}.M_X(t)$$

وهذا يعني ان العزم المركزي ذو المرتبة r هو :

$$E(X-\mu)^r = M^{(r)}_{(X-\mu)}(0), r = 1, 2, 3, \ldots$$

مثال (٢٩) : اذا علمت ان الدالة المولدة لعزوم X حول نقطة الاصل هي $M_X(t) = e^{2t+8t^2}$. جد الدالة المولدة للعزوم المركزية ثم جد العزم المركزي الثالث .

لحل : نجد اولاً الوسط الحـ X

$$M'_X(t) = e^{2t+8t^2}.(2+16t). \quad \therefore \mu_x = M'_X(0) = 2$$

فاذن

$$M_{(X-2)}(t) = e^{-2t}.e^{2t+8t^2} = e^{8t^2}$$

عليه فان :

$$M'_{(X-2)}(t) = e^{8t^2}.(16t) = 16te^{8t^2}$$

$$M''_{(X-2)}(t) = 16(e^{8t^2} + te^{8t^2}(16t)) = 16e^{8t^2}(1+16t^2)$$

$$M'''_{(X-2)}(t) = 16(e^{8t^2}(32t) + (1+16t^2)e^{8t^2}(16t))$$

$$\therefore E(X-2)^3 = M'''_{(X-2)}(0) = 0$$

٢ ـ ٢ ـ ٤ : الدالة المولدة للعزوم المطلقة المركزية central absolute moment generating function

اذا تم اختيار الدالة $g(x) = |x-\mu_x|$ عندئذٍ يمكن توليد العزم المركزي المطلق ذي المرتبة r من خلال التعويض عن $g(x)$ بـ $|x-\mu_x|$ في (*) .

وهذا يعني ان

$$M_{|X-\mu|}(t) = Ee^{t|X-\mu|}$$

وان

$$E|X-\mu|^r = M^{(r)}_{|X-\mu|}(0), r = 1, 2, 3, \ldots$$

مثال (٣٠) : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = e^{-x}, x \geq 0$. جد الدالة المولدة للعزوم المطلق المركزية ثم احسب العزم المطلق الاولى والثاني .

الحل : ان الوسط لقيم X في هذا التوزيع هو $\mu = 1$ فاذن

$$M_{|X-1|}(t) = \int_0^{\infty} e^{t|x-1|} \cdot e^{-x} dx$$

$$= \int_0^{1} e^{t|x-1|} \cdot e^{-x} dx + \int_1^{\infty} e^{t|x-1|} \cdot e^{-x} dx$$

وحيث ان $x - 1 < 0$ في الفترة $[0, 1]$ فاذن $|x-1| = -(x-1) = 1 - x$ وان $x - 1 > 0$ في الفترة $[1, \infty[$ فاذن $|x - 1| = x - 1$. عليه فان

$$M_{|X-1|}(t) = \int_0^{1} e^{t(1-x)} \cdot e^{-x} dx + \int_1^{\infty} e^{t(x-1)} \cdot e^{-x} dx$$

$$= e^{t} \int_0^{1} e^{-x(1+t)} dx + e^{-t} \int_1^{\infty} e^{-x(1-t)} dx$$

$$= \frac{e^{t}}{1+t} + \frac{2te^{-1}}{1-t^2}$$

لاحظ ان $M_{|X-1|}(0) = 1$ فاذن

$$M'_{|X-1|}(t) = \frac{te^{t}}{(1+t)^2} + \frac{2e^{-1}(1+t^2)}{(1-t^2)^2}$$

عليه فان

$$E|X-1| = M'_{|X-1|}(0) = 2e^{-1}$$

كذلك فان

$$M''_{|X-1|}(t) = \frac{e^{t}(1+t^2)}{(1+t)^3} + \frac{4e^{-1}t(3+t^2)}{(1-t^2)^3}$$

فاذن :

$$E|X-1|^2 = E(X-1)^2 = V(X) = M''_{(X-1)}(0) = 1$$

ويطلب من القاريء حساب قيمة العزم المطلق الثالث والرابع .

٢ ـ ٢ ـ ٥ : الدالة المولدة للعزوم العاملية

Factorial moment generating function

في حالة اختيار الدالة $g(x) = \frac{x \ln t}{t}, t \neq 0$ عندئذٍ يمكن توليد العزم العاملي ذي المرتبة r من خلال التعويض عن $g(x)$ بـ $\frac{x \ln t}{t}$ في (*)
فاذا رمزنا للدالة المولدة للعزوم العاملية بالرمز $M(t)$ فذلك يعني ان :

$$M(t) = Ee^{t \cdot g(x)} = Ee^{t \cdot \frac{X \ln t}{t}}, \quad M(1) = 1$$
$$= Ee^{X \ln t} = Ee^{\ln t^X} = Et^X = M_X(\ln t)$$

وباستخدام سلسلة مكلورين (ووفق الصيغة *) يمكن البيان ان :

$$M(t) = M_X(\ln t) = 1 + (\ln t)EX + \frac{(\ln t)^2}{2!}EX^2 + \frac{(\ln t)^3}{3!}EX^3 + \dots$$

ومن خلال هذه السلسلة يمكن توضيح عملية توليد العزوم العاملية للتوزيع وعلى النحو الآتي :
بايجاد المشتقة الاولى للدالة $M(t)$ نسبة الى t نحصل على :

$$M'(t) = \frac{1}{t}EX + \frac{\ln t}{t}EX^2 + \frac{(\ln t)^2}{2!t}EX^3 + O'(\ln t)$$

حيث $O'(\ln t)$ تشير الى حدود لاحقة تمثل مشتقات من المرتبة الاولى تتضمن $\ln t$ بقوى عليا . وبجعل $t = 1$ نحصل على $M'(1) = EX$ وبايجاد المشتقة الثانية للدالة $M(t)$ نسبة الى t نحصل على :

$$M''(t) = -\frac{1}{t^2}EX + \frac{1-\ln t}{t^2}EX^2 + \frac{2\ln t - (\ln t)^2}{2t^2}EX^3 + O''(\ln t)$$

حيث$O''(\ln t)$ تشير الى حدود لاحقة تمثل مشتقات من المرتبة الثانية تتضمن $\ln t$ بقوى عليا . وبجعل $t = 1$ انحصل على :

$$M''(1) = -EX + EX^2 = EX^2 - EX = EX(X-1)$$

ووفق نفس الاجراء الموضح اعلاه يمكن البيان ايضاً ان :

$$M'''(1) = EX^3 - 3EX^2 + 2EX = EX(X-1)(X-2)$$

وهذا يعني ان العزم العاملي ذو المرتبة r هو

$$M^{(r)}(1) = E \prod_{j=1}^{r} (X - j + 1), r = 1, 2, 3 \ldots$$

مثال (٣١) : افرض أن $M_X(t) = e^{\frac{1}{2}t^2}$ تمثل الدالة المولدة لعزوم X حول نقطة الاصل . جد الدالة المولدة للعزوم العاملية ثم جد العزم العاملي الاول والثاني .

الحل :

$$M(t) = M_X(\ln t) = e^{\frac{1}{2}(\ln t)^2}$$

لاحظ ان $M(1) = 1$. الان

$$M'(t) = e^{\frac{1}{2}(\ln t)^2} \cdot \frac{\ln t}{t} = M(t) \cdot \frac{\ln t}{t}$$

فاذن $M'(1) = EX = 0$

كذلك فان

$$M''(t) = M(t) \cdot \left(\frac{1 - \ln t + (\ln t)^2}{t^2} \right)$$

فاذن

$$M''(1) = EX(X-1) = 1$$

ويطلب من القاريء حساب العزم العاملي الثالث والرابع والخامس

٢ ـ ٢ ـ ٦ : الدالة المولدة الاحتمالية

Probability generating function

ان مفهوم هذه الدالة المولدة يقترن بحالة المتغيرات العشوائية المتقطعة وهي لاتختلف بشيء عن الدالة المولدة للعزوم العاملية سوى انها ممكنة الاستخدام لتوليد عزوم التوزيعات المتقطعة كأسلوب بديل للدالة المولدة للعزوم حول نقطة الاصل. وعلى فرض ان $a_0, a_1, a_2, \ldots$ تمثل سلسلة من الاعداد الحقيقية وان $A(t)$ دالة بدلالة t بحيث ان

$$A(t) = \sum_{x=0}^{x} a^k \cdot t^k, \; -h < t < h \,.\, h > 0$$

عندئذٍ يقال ان $A(t)$ هي دالة مولدة للسلسلة $\{a_k\}$ اذا كانت $A(t)$ متقاربة لجميع قيم t المعرفة في الفترة $[-h, h]$. واذا كان X متغيراً عشوائياً غير سالب يسلك وفق دالة كتلة احتمالية $P(x = k)$ عندئذٍ اذا تم اختيار $a_k = P(X=k)$ و $k = 0, 1, 2, \ldots$ فان :

$$A(t) = \sum_{x=0}^{x} P(x) t^x = E(t^X)$$

وفي هذه الحالة تسمى الدالة $A(t)$ بالدالة المولدة الاحتمالية للمتغير X. ويلاحظ ان $A(1) = 1$ طالما ان $P(x)$ هي دالة كتلة احتمالية. وهذا يعني ان $A(t)$ متقاربة لجميع قيم t المعرفة في الفترة $[-1, 1]$. ويمكن الملاحظة وبسهولة العلاقة التي تربط ما بين $M_X(t), A(t)$ وهي :

$$M_X(t) = Ee^{tX} = E(e^t)^X = A(e^t)$$

كذلك فان

$$\left.\frac{d^r A(t)}{dt^r}\right|_{t=1} = A^{(r)}(1) = M^{(r)}(1) = E \prod_{j=1}^{r} (X - j + 1)$$

مثال (٣٢) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{3^x e^{-3}}{x!}, \; x = 0, 1, 2, \ldots$ جد الدالة المولدة الاحتمالية .

الحل :

$$A(t) = \sum_{x=0}^{x} t^x \cdot \frac{3^x e^{-3}}{x!} = e^{-3} \sum_{x=0}^{x} \frac{(3t)^x}{x!}$$
$$= e^{-3} \cdot e^{3t} = e^{3(t-1)}$$

لاحظ ان $A(1) = 1$. كذلك فان

$$M_X(t) = A(e^t) = e^{3(e^t - 1)}$$

وهي الدالة المولدة للعزوم حول نقطة الاصل التي حصلنا عليها في المثال (١) من الفقرة (٢ ــ ٢ ــ ١) . كذلك فان

$$A'(t) = e^{3(t-1)} \cdot 3 = 3A(t) \qquad \therefore A'(1) = 3 = EX$$
$$A''(t) = e^{3(t-1)} \cdot 9 = 9A(t) \qquad \therefore A''(1) = 9 = EX(X-1)$$
$$A'''(t) = e^{3(t-1)} \cdot 27 = 27A(t) \qquad \therefore A'''(1) = 27 = EX(X-1)(X-2)$$

وفي هذه الحالة يمكن ملاحظة ان $A^{(r)}(1) = 3^r , r = 1, 2, \ldots$

٢ ــ ٣ : الدالة المميزة Characteristic function

وتسمى في بعض الاحيان « الدالة الوصفية » التي تعد بحق من اهم دوال توليد العزوم لما تتمتع به من خصائص تطبيقية جعلتها تقف في مقدمة هذا النوع من الدوال . وكما لاحظنا لدى دراستنا لموضوع الدالة المولدة للعزوم حول نقطة الاصل فان بعض التوزيعات الاحتمالية قد لاتمتلك دالة مولدة للعزوم بسبب عدم تحقق خاصية التقارب المطلق مما يسبب ذلك عدم امكانية التعرف على عزوم ذلك التوزيع وخصوصاً مايتعلق الامر بالوسط والتباين . في حين وكما سنلاحظ لاحقاً فان كل توزيع احتمالي يمتلك دالة مميزة . ان البراهين التي سترد في هذه الفقرة سوف تخصص لحالة المتغيرات العشوائية المستمرة والاسلوب ذاته ينطبق على حالة المتغيرات العشوائية المتقطعة بمجرد استبدال رمز التكامل برمز الجمع .

بفرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x)$ معرفة قيمه على فضاء عينه مثل Ω ، وان t متغير آخر و h عدد موجب بحيث ان $-h < t < h$ وان $i = \sqrt{-1}$ ، عندئذٍ تعرف الدالة المميزة ، التي غالباً ما يرمز لها بالرمز $\phi(t)$ ، على النحو التالي :

$$\phi(t) = Ee^{itX} = E(\cos tX + i \sin tX)$$

ان هذا التعريف ناتج عن مايلي : يمكن ملاحظة ان مفكوك سلسلة مكلورين الى e^{itX} هو :

$$e^{itX} = \sum_{k=0}^{\infty} \frac{(itX)^k}{k!} = 1 + itX + \frac{(itX)^2}{2!} + \frac{(itX)^3}{3!} + \frac{(itX)^4}{4!} + \ldots$$

$$= \left(1 - \frac{(tX)^2}{2!} + \frac{(tX)^4}{4!} - \frac{(tX)^6}{6!} + \ldots\right)$$

$$+ i\left(tX - \frac{(tX)^3}{3!} + \frac{(tX)^5}{5!} - \frac{(tX)^7}{7!} + \ldots\right)$$

$$= \cos tX + i \sin tX$$

كذلك فان

$$\phi(t) = Ee^{itX} = \int_{\Omega} e^{itx} . f(x)\, dx = M_X(it)$$

٢ ـ ٣ ـ ١ : خصائص الدالة المميزة :

فيما يلي بعض الخصائص التي تتمتع بها الدالة المميزة وهي :

١ ـ ان $\phi(0) = 1$ وذلك بسبب ان $\sin(0) = 0, \cos(0) = 1$

٢ ـ ان $|\phi(t)| \leq 1$. اي ان $\phi(t)$ دالة محدودة لجميع قيم t .

البرهان :

$$|\phi(t)| = \left|\int_{\Omega} e^{itx}.f(x)dx\right| \le \int_{\Omega} |e^{itx}|.f(x)dx, f(x) > 0$$

لكن

$$|e^{itx}| = |\cos tx + i\sin tx| = \sqrt{\cos^2 tx + \sin^2 tx} = 1$$

فاذن

$$|\phi(t)| \le \int_{\Omega} f(x)dx = 1$$

من هذه الخاصية نلاحظ ان التكامل متقارب دائما الامر الذي يستدعي القول بان الدالة المميزة موجودة دائماً مهما كان التوزيع الاحتمالي الى X .

٣ ـ ان $\phi(t)$ دالة مستمرة ومنتظمة بدلالة t في الفترة $(-\infty, \infty)$.

البرهان : افرض ان $h \neq 0$ عدد حقيقي . ان المطلوب برهانه في هذه الخاصية هو ان $\lim_{h \to 0} \phi(t+h) = \phi(t)$ لجميع قيم t .

الآن : $\phi(t+h) = Ee^{i(t+h)X}$

وان

$$|\phi(t+h) - \phi(t)| = \left|\int_{\Omega} (e^{i(t+h)x} - e^{itx})f(x)dx\right|$$

$$\le \int_{\Omega} |e^{i(t+h)x} - e^{itx}|f(x)dx$$

$$= \int_{\Omega} |e^{itx}(e^{ihx} - 1)|f(x)dx = \int_{\Omega} |e^{itx}|.|e^{ith} - 1|.f(x)dx$$

$$= \int_{\Omega} |e^{ihx} - 1|f(x)dx \; ; |e^{itx}| = \sqrt{\cos^2 tx + \sin^2 tx} = 1.$$

لكن

$$|e^{ihx} - 1| \le |e^{ihx}| + |-1| = 2 \; ; |e^{ihx}| = 1$$

فاذن

$$\int_{\Omega} |e^{ihx} - 1| f(x)\,dx \le 2 \int_{\Omega} f(x)\,dx = 2$$

وهذا يعني ان

$$|\phi(t+h) - \phi(t)| \le 2$$

فاذن $\phi(t)$ دالة محدودة.

كذلك فان

$$\lim_{h \to 0} \left| \phi(t+h) - \phi(t) \right| \le \int_{\Omega} \lim_{h \to 0} |e^{ihx} - 1| f(x)\,dx = 0$$

وهذا يعني ان

$$\lim_{h \to 0} \phi(t+h) = \phi(t)$$

٤ ـ ان $\phi(t)$ و $\phi(-t)$ دالتان مترافقتان(*) conjugate functions

البرهان :

$$\phi(t) = Ee^{itX} = E(\cos tX + \sin tX)$$

فاذن

$$\overline{\phi(t)} = E(\cos tX - i \sin tX) = E(\cos(-t)X + i \sin(-t)X)$$
$$= Ee^{-itX} = \phi(-t).$$

٥ ـ ان $\phi(t)$ هي دالة وحيدة . اي المقصود من ذلك هو ان لكل توزيع احتمالي هنالك دالة مميزة واحدة فقط والعكس صحيح اي ان لكل دالة مميزة هنالك توزيع احتمالي واحد لمتغير عشوائي مثل X . ان برهان هذه الخاصية تتم من خلال « نظرية الانعكاس لـ فورايز "Fourier's inversion theorem التي تنص بما يلي : لتكن $f(x)$ دالة الكثافة الاحتمالية الى X وان $\phi(t)$ تمثل الدالة المميزة . عندئذٍ يمكن استنتاج الدالة $f(x)$ اذا علمت $\phi(t)$ والعكس صحيح اي يمكن ايجاد $\phi(t)$ اذا علمت $f(x)$. ونظرا لكون برهان هذه النظرية يقع خارج

(*) ، ليكن a, b عددين حقيقيين وان $i = \sqrt{-1}$ يسمى كل من $Z_1 = a + ib$, $Z_2 = a - ib$ عدد معقد Complex number وان Z_1 مرافق للعدد Z_2 وان Z_2 مرافق للعدد Z_1 . ويصطلح لذلك بالشكل $\bar{Z}_2 = Z_1$ $\bar{Z}_1 = Z_2$ اي ان $\bar{Z}_1 = \overline{(a+ib)} = a - ib = Z_2$.

نطاق هذا الكتاب عليه سوف نكتفي بالتعامل مع هذه النظرية من الناحية التطبيقية . ان النتائج المستخلصة من هذه النظرية هي :

١ ـ ان

$$\phi(t) = \int_{\Omega} e^{itx} \cdot f(x)\, dx$$

٢ ـ ان

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \cdot \phi(t)\, dt$$

٣ ـ فيما يخص الاستنتاج الثاني هنالك مشكلة تحديد نوع التوزيع الاحتمالي (متقطع أم مستمر) . ان ذلك يمكن تحديده وفق الاختبار التالي :

افرض الدالة التالية :

$$L_c = \int_{-c}^{c} e^{-itx} \cdot \phi(t)\, dt$$

حيث ان الفترة [c , c −] هي مجموعة جزئية من R . فاذا لاحظنا ان $\lim_{c \to \infty} \frac{L_c}{2c} = 0$ فذلك يعني ان f (x) مستمره . وفي غير ذلك نقول ان f (x) هي دالة لتوزيع احتمالي متقطع . وسوف نوضح ذلك من خلال الامثلة التي سنوردها في الفقرة (٢ ـ ٣ ـ ٢) .

نأتي الآن الى توضيح عملية توليد عزوم توزيع احتمالي حول نقطة الاصل باستخدام الدالة المميزة . ان مفكوك سلسلة مكلورين الى e^{itX} هو :

$$e^{itX} = \sum_{k=0}^{\infty} \frac{(itX)^k}{k!} = 1 + itX + \frac{(itX)^2}{2!} + \frac{(itX)^3}{3!} + \ldots$$

وبأخذ توقع الطرفين نحصل على :

$$\phi(t) = Ee^{itX} = 1 + itEX + \frac{i^2t^2}{2!} EX^2 + \frac{i^3t^3}{3!} EX^3 + \ldots$$

الان بايجاد المشتقة الاولى الى $\phi(t)$ نسبة الى t نحصل على

$$\phi'(t) = iEX + i^2tEX^2 + \frac{i^3t^2}{2!}EX^3 + O'(t)$$

حيث $O'(t)$ تعني حدود لاحقه تمثل مشتقات من المرتبة الاولى تتضمن t بقوى عليا . وبجعل $t = 0$ في $\phi'(t)$ نحصل على $\phi'(0) = i.EX$ وبايجاد المشتقة الثانية الى $\phi(t)$ نسبة الى t وجعل $t = 0$ نحصل على $\phi''(0) = i^2EX^2$. ووفق هذا السياق يمكن الاستنتاج بأن

$$\phi^{(r)}(0) = i^rEX^r$$

وحيث ان $EX^r = M_X^{(R)}(0)$ فاذن

فاذن

$$\phi^{(r)}(0) = i^rEX^r = i^rM_X^{(r)}(0)$$
$$EX^r = i^{-r}.\phi^{(r)}(0) = M_X^{(r)}(0)$$

والشكل الاخير يمثل العلاقة بين الدالة المولدة للعزوم حول نقطة الاصل والدالة المميزة .

٢ ـ ٣ ـ ٢ : تطبيقات :

فيما يلي بعض الامثلة التوضيحية عن الدالة المميزة .

مثال (٣٣) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $P(x) = \frac{5^xe^{-5}}{x!}$ ، $x = 0, 1, 2, \ldots$. جد الدالة المميزة .

الحل :

$$\phi(t) = \sum_{x=0}^{\infty} e^{itx}.\frac{5^xe^{-5}}{x!} = e^{-5}\sum_{x=0}^{\infty}\frac{(5e^{it})^x}{x!}$$

$$= e^{-5}.e^{5e^{it}} = e^{5(e^{it}-1)}$$

مثال (٣٤) : افرض ان X متغير عشوائي بدالة كثافة احتمالية , $f(x) = e^{-x}$
$0 \geq x$ جد الدالة المميزة ، ثم جد العزمين الاول والثاني حول نقطة الاصل

الحل :

$$\phi(t) = \int_0^{\infty} e^{itx} . e^{-x} dx = \int_0^{\infty} e^{-x(1-it)} dx = \frac{1}{1-it}, t \neq \frac{1}{i}$$

كذلك فان

$$\phi'(t) = \frac{i}{(1-it)^2} \quad \therefore \phi'(0) = i$$

وهذا يعني ان

$$EX = \frac{1}{i} \phi'(0) = \frac{1}{i} . i = 1$$

وان

$$\phi''(t) \frac{2i^2}{(1-it)^3} \quad \therefore \phi''(0) = 2i^2$$

فاذن

$$EX^2 = \frac{1}{i^2} \phi''(0) = 2$$

مثال (٣٥) : لتكن $\phi(t) = (a + be^{it})^n$ ، حيث $n ; a + b = 1 ; 0 < a, b < 1$ عدد موجب صحيح ، تمثل الدالة المميزة الى X . جد التوزيع الاحتمالي الى X .

الحل : باستخدام نظرية ثنائي الحدين يمكن ملاحظة ان :

$$\phi(t) = (a + be^{it})^n = \sum_{k=0}^{n} C_k^n (be^{it})^k . (a)^{n-k}$$

$$= \sum_{k=0}^{n} C_k^n (b)^k . (a)^{n-k} . e^{itk}$$

الان :

$$L_c = \int_{-c}^{c} e^{-itx} . \phi(t) dt = \int_{-c}^{c} e^{-itx} . \sum_{k=0}^{n} C_k^n (b)^k . (a)^{n-k} . e^{itk} . dt$$

$$= \sum_{k=0}^{n} C_k^n (b)^k . (a)^{n-k} \int_{-c}^{c} e^{-it(x-k)} dt$$

ونلاحظ مايلي : اذا كان $x \neq k$ عندئذٍ

$$L_c = \sum_{k=0}^{n} C_k^n (b)^k . (a)^{n-k} . \left[\frac{e^{-it(x-k)}}{-i(x-k)} \right]_{-c}^{c}$$

$$= \sum_{k=0}^{n} c_k^n (b)^k . (a)^{n-k} \left[\frac{e^{ic(x-k)} - e^{-ic(x-k)}}{i(x-k)} \right]$$

لكن

$\sin Z = \frac{1}{2i} (e^{iZ} - e^{-iZ})$ ، وبفرض ان $Z = c(X - k)$. عندئذٍ

$$2i \sin [c(x-k)] = e^{ic(x-k)} - e^{-ic(x-k)}$$

وهذا يعني ان

$$\frac{L_c}{2c} = \sum_{k=0}^{n} c_k^n (b)^k . (a)^{n-k} . \left[\frac{\sin(c(x-k))}{c(x-k)} \right]$$

فاذن

$$\lim_{c \to \infty} \frac{L_c}{2c} = \sum_{k=0}^{n} c_k^n (b)^k (a)^{n-k} . \lim_{c \to \infty} \left[\frac{\sin(c(x-k))}{c(x-k)} \right] = 0$$

وهذا يعني انه عندما $X \neq k$ هنالك انقطاع **discontinuity** في الدالة لكن اذا كان $X = k$ عندئذٍ

$$L_c = \sum_{k=0}^{n} c_k^n (b)^k . (a)^{n-k} \int_{-c}^{c} dt = 2c \sum_{k=0}^{n} c_k^n (b)^k . (a)^{n-k}$$

لكن وحسب نظرية ثنائي الحدين فان

$$\sum_{k=0}^{n} c_k^n b^k . a^{n-k} = (a + b)^n = 1$$

فاذن

$$L_c = 2c \rightarrow \frac{L_c}{2c} = 1 \qquad \therefore \lim_{c \rightarrow \infty} \frac{L_c}{2c} = 1$$

وهذا يعني ان الدالة غير مستمرة عند $X = k$ نستنتج من ذلك ان التوزيع الاحتمالي الى X هو توزيع متقطع . وحيث ان $\sum_{k=0}^{n} c_k^n b^k a^{n-k} = 1$ فذلك يعني ان دالة الكتلة الاحتمالية الى X هي :

$$P(X = K) = c_k^n b^k . a^{n-k} \quad , k = 0, 1, 2, \ldots, n$$

وسوف نلاحظ في الفقرة (٥ ــ ٣) ان هذه الدالة ماهي الا دالة توزيع ثنائي الحدين .

مثال (٣٦) : افرض ان $\phi(t) = e^{-\frac{1}{2}t^2}$ تمثل الدالة المميزة لمتغير عشوائي X. جد التوزيع الاحتمالي الى X.

الحل :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} . e^{-\frac{1}{2}t^2} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(t^2 + 2itx)} dt$$

وباكمال المربع داخل القوس في التكامل الاخير نحصل على :

$$f(x) = \frac{1}{2\pi} e^{-\frac{1}{2}x^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(t+ix)^2} dt$$

وبفرض ان $Z = t + ix$ فان $dt = dz$. عليه فان .

$$f(x) = \frac{1}{2\pi} e^{-\frac{1}{2}x^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz$$

لكن $\sqrt{2\pi} = \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}\, dz$، (لاحظ برهان ذلك في الفقرة (٦ ــ ٢)، فاذن

$$f(x) = \frac{1}{2\pi} e^{-\frac{1}{2}x^2} \cdot \sqrt{2\pi} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

وان

$$\frac{Lc}{2c} = \frac{1}{2c} \cdot \frac{1}{2\pi} e^{-\frac{1}{2}x^2} \int_{-c}^{c} e^{-\frac{1}{2}z^2}\, dz$$

فاذن

$$\lim_{c\to\infty} \frac{Lc}{2c} = \frac{1}{4\pi} e^{-\frac{1}{2}x^2} \cdot \lim_{c\to\infty} \left(\frac{1}{c}\right) \cdot \lim_{c\to\infty} \int_{-c}^{c} e^{-\frac{1}{2}z^2}\, dz = 0$$

وهذا يعني ان الدالة $f(x)$ مستمرة عند اية قيمة الى X المعرفة في الفترة $(-\infty,\infty)$ وانها محدودة طالما أن $\frac{1}{\sqrt{2\pi}} \int_{-c}^{c} e^{-\frac{1}{2}x^2}\, dx \le \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\, dx = 1$ ، عليه نستنتج

ان دالة التوزيع الاحتمالي للمتغير X هي :

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty < x < \infty$$

وسوف نلاحظ ان هذه الدالة تمثل دالة التوزيع الطبيعي المعياري الذي سيأتي ذكره في الفقرة (٦ ــ ٢ ــ ٧) .

٢ ـ ١ : لكل حالة من الحالات التالية بين فيما اذا كان توقع الدالة $g(x)$ موجوداً ام غير موجود مع ذكر السبب .

أ ـ $p(x) = \frac{6}{\pi^2 x^2}; x = 1, 2, 3, \ldots \qquad , g(x) = X^2 , g(x) = X$

ب ـ $p(x) = (0.3)(0.7)^x, x = 0, 1, 2, \ldots \quad , g(x) = (0.3)^x, g(x) = 4^x$

جـ ـ $f(x) = be^{-bx}, x \geq 0, b > 0 \qquad g(x) = e^x, g(x) = e^{bX}, g(x) = \frac{1}{X}$

د ـ $f(x) = \frac{1}{\pi(1+x^2)}, -\infty < x < \infty, g(x) = X^2 + 1, g(X) = \pi X$

٢ ـ ٢ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = \frac{X+c}{18}, -2 < x < 4$
جد قيمة الثابت C ثم احسب مايلي :

$$EX, E(X+4)^2, E(X-3)^2, E(X-EX)^3, E(3X - E(X+2)^2)^2, V(X)$$

٢ ـ ٣ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = 2x, 0 < x < 1$ جد مايلي $E\sqrt{X}, E\frac{1}{X}, E\ln X$ ثم تحقق مما يلي :

$$E \ln X = \ln EX \quad , E\frac{1}{X} = \frac{1}{EX} \quad , E\sqrt{X} = \sqrt{EX}$$

٢ ـ ٤ : لتكن $f(x) = \frac{c}{x^5}, x \geq 1$ تمثل دالة الكثافة الاحتمالية الى X . جد :

أ ـ قيمة الثابت C

ب ـ العزوم اللامركزية الثلاثة الاولى حول النقطة 2 .

جـ ـ العزوم الثلاثة الاولى حول نقطة الاصل .

د ـ العزوم المركزية الثلاثة الاولى .

هـ ـ العزوم العاملية الثلاثة الاولى .

و ـ العزم المطلق الثالث .

٢ ـ ٥ : لمعطيات السؤال (٢ ـ ٣) جد ماهو مطلوب في السؤال (٢ ـ ٤) .

٢ ـ ٦ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = 6x(1-x), 0 < x < 1$ جد مايلي :

أ ـ الوسط والتباين في هذا التوزيع .

ب ـ العزم ذا المرتبة r حول نقطة الاصل .

جـ ـ الانحراف المطلق .

د ـ العزم العاملي الثالث والرابع .

هـ ـ الوسط التوافقي .

٢ ـ ٧ : لتكن $f(x)=cx^2, 0< x< 1$ تمثل دالة الكثافة الاحتمالية للمتغير X .

جد قيمة C ثم جد مايلي :

أ ـ الوسط والتباين الى X ثم احسب $P_r(\mu - \sigma < X < \mu + \sigma)$

ب ـ الوسط التوافقي

جـ ـ الانحراف المطلق .

د ـ العزم ذا المرتبة r حول نقطة الاصل .

هـ ـ العزم العاملي الثالث .

و ـ العزم المركزي الثالث .

٢ ـ ٨ : اذا علمت ان دالة الكتلة الاحتمالية $p(x)$ لمتغير عشوائي X موصوفة بالشكل التالي :

x :	1	2	3
p(x) :	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

جد مايلي :

أ ـ الوسط والتباين الى X

ب ـ الوسط التوافقي

جـ ـ العزم ذا رتبة r حول نقطة الاصل .

د ـ العزم العاملي الثالث .

٢ ـ ٩ : افرض ان دالة الكتلة الاحتمالية للمتغير X موصوفة بالشكل التالي :

x :	−3	−2	−1	0	1	2	3
p(x) :	0·05	0·10	0·25	0·10	0·25	0·15	0·10

جد مايلي :

أ ـ الوسط والتباين الى X .

ب ـ الوسط والتباين الى $Y = 2X - 3$, $Y = X^2 + 4X$.

جـ ـ الوسط التوافقي الى X .

د ـ العزم الثالث الى X حول نقطة الاصل .

هـ ـ العزم العاملي الثالث .

٢ ـ ١٠ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x), 0 < x < b$

برهن ان $EX = \int_0^b (1 - F(x))\,dx$, $EX^2 = 2\int_0^b x(1 - F(x))\,dx$

واذا علمت ان $f(x) = e^{-x}, x \geq 0$. جد $V(X), EX$ وفق هاتين الصيغتين .

٢ ـ ١١ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = \frac{1}{4}, 0 < x < 4$

أ ـ جد الدالة المولده لعزوم X حول نقطة الاصل .

ب ـ تحقق من ان $M_x(0) = 1$

جـ ـ جد الوسط والتباين الى X باستخدام $M_X(t)$

د ـ جد الدالة المولدة لعزوم $Y = 2X + 4$ حول نقطة الاصل .

هـ ـ جد الدالة المولدة لعزوم الدرجة المعيارية في هذا التوزيع .

و ـ جد الدالة المميزة لهذا التوزيع .

٢ ـ ١٢ : اذا علمت ان العزم ذا المرتبة r حول نقطة الاصل في توزيع احتمالي معين هو $EX^r = 2^r(r+1)!$ جد الدالة المولده لعزوم هذا التوزيع حول نقطة الاصل ثم جد الدالة المولدة التراكمية . ماهي قيمة $V(X), EX$.

٢ ـ ١٣ : افرض ان العزم ذا المرتبة r حول نقطة الاصل في توزيع احتمالي معين هو $EX^r=(2r+1)^{-1}$ جد الدالة المولده لعزوم التوزيع حول نقطة الاصل .

٢ ـ ١٤ : ليكن X متغير عشوائي بدالة كتلة احتمالية $p(x)=2^{-x}, x=1, 2, \ldots$ جد :

أ ـ الدالة المولدة لعزوم X حول نقطة الاصل . ثم جد $V(X), EX$

ب ـ الدالة المولدة لعزوم X حول النقطة 4 ثم جد قيمة العزم الثالث .

جـ ـ الدالة المولدة للعزوم المركزية ثم جد قيمة العزم الثالث .

د ـ الدالة المولدة للعزوم العاملية مع حساب قيمة العزم الرابع .

هـ ـ الدالة المولدة للعزوم المطلقة ثم جد قيمة الانحراف المطلق .

و ـ الدالة المولدة الاحتمالية .

٢ ـ ١٥ : جد الدالة المميزة لكل توزيع من التوزيعات التالية مع حساب الوسط والتباين

$$f(x) = ae^{-ax}, a > 0, x \geq 0 \quad , f(x) = \frac{1}{2} e^{-|x|}, -\infty < x < \infty$$

$$p(x) = \frac{5^x e^{-5}}{x!}; x = 0, 1, 2, \ldots , p(x) = \frac{1}{6} \quad , x = 1, 2, \ldots, 6 .$$

٢ ـ ١٦ : اذا علمت ان الدالة المميزة لتوزيع احتمالي معين هي $\phi(t) = e^{2it-4t^2}$
جد التوزيع الاحتمالي للمتغير X ثم جد الوسط والتباين من خلال $\phi(t)$.

٢ ـ ١٧ : اذا علمت ان $EX^r = r!$ ، $r = 1, 2, \ldots$. جد مايلي :

أ ـ الدالة المولدة لعزوم X حول نقطة الاصل .

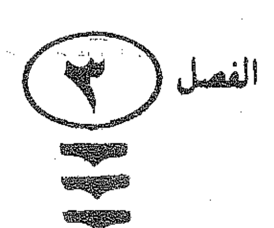
ب ـ الدالة المميزة .

جـ ـ هل يمكن القول ان الدالة المميزة التي حصلت عليها في (ب) مشتقة من توزيع احتمالي معرف بالدالة $f(x) = e^{-x}; x \geq 0$.

٢ ـ ١٨ : افرض ان $\phi(t)$ تمثل الدالة المميزة لتوزيع احتمالي لمتغير عشوائي . جد الدالة المميزة الى : أ ـ $Y = a + bX$. ب ـ الدرجة المعيارية Z

٢ ــ ١٩ : لتكن $\phi(t)$ تمثل الدالة المميزة لتوزيع احتمالي لمتغير عشوائي . وافرض ان $\psi(t) = \ln \phi(t)$. برهن ان $EX = \frac{\psi'(0)}{i}, V(X) = \frac{\psi''(0)}{i^2}$

٢ ــ ٢٠ : اذا علمت ان $M(t)$ تمثل الدالة المولدة للعزوم العاملية لتوزيع احتمالي معين وان $\psi(t) = \ln M(t)$. برهن ان $V(X) = \psi''(1) - \psi'(1)$





مقاييس اخرى عن التوزيعات الاحتمالية

الفصل الثالث

مقاييس اخرى عن التوزيعات الاحتمالية

لاحظنا في الفصل السابق كيفية حساب بعض المؤشرات (المقاييس) الاحصائية كالوسط والتباين وغيرها كتطبيقات لمفهوم التوقع الرياضي ودوال توليد العزوم. في هذا الفصل سوف نتطرق لدراسة بعض المقاييس الاخرى الممكنة الحصول من توزيع احتمالي بعضها يمثل مقاييس موقع (نزعة مركزية) والبعض الآخر يمثل مقاييس تشتت بالاضافة الى ذلك سوف نستعرض اهم المقاييس التي تصف هيئة وشكل التوزيع الاحتمالي . كذلك سوف نتطرق لمفهوم التوزيعات الاحتمالية المقطوعة (المبتورة) والهدف من دراستها .

٣ ــ ١ : المنوال Mode .

يعد المنوال احد مقاييس النزعة المركزية (مقياس موقعي) قيمته تكون معرفة على المحور السيني . ويعرف المنوال بأنه قيمة من قيم X المعرفة في Ω التي تجعل f (x) في نهايتها العظمى اما في حالة المتغيرات المتقطعة فان المنوال يمثل قيمة حقيقة (قد تكون معرفة في Ω او قد لاتكون) تجعل p (x) اكبر ما يمكن . ان وحدات قياس المنوال هي نفس وحدات قياس المتغير X . ان الهدف من دراسة المنوال هو تكوين فكرة عن القيمة العظمى للدالة f (x) او p (x) اضافة الى كونه مقياس بديل للمتوسط في حالة عدم امكانية ايجاد الاخير . وفي حالة المتغيرات العشوائية المستمرة . ووفق تعريف المنوال . فانه يمكن الحصول على هذا المقياس (اذا كان موجود) من خلال البحث عن قيمة (او قيم X') التي تجعل f (x) في نهايتها العظمى . وذلك يعني البحث عن قيمة (او قيم X) التي تحقق المعادلة التفاضلية :

$$f'(x) = \frac{df(x)}{dx} = 0 \quad \text{بشرط ان} \quad f''(x) = \frac{d^2 f(x)}{dx^2} < 0 \quad \text{عند}$$

قيمة (او قيم) X التي تحقق $f'(x) = 0$. اما في حالة المتغيرات العشوائية المتقطعة فان المنوال (اذا كان موجود) يمثل قيمة معرفة على المحور السيني التي تحقق المتباينة $\dots < P(x-2) < p(x-1) < p(x) > p(x+1) > p(x+2) > \dots$ واذا لوحظ ومن خلال هذه المتباينة ان $\dots < P(x-1) < P(x) = P(x+1) > P(x+2) > \dots$ عندئذ يقال ان التوزيع الاحتمالي يمتلك منوالين معرفين بالقيمتين $x, x+1$. اما اذا كانت جميع الكتل الاحتمالية متساوية القيمة عند ذاك يقال ان التوزيع الاحتمالي للمتغير العشوائي X لا يمتلك منوال . وخير مثال على ذلك، هو التوزيع المنتظم المتقطع الموضح بالمثال (٤) لاحقاً .

مثال (١) : افرض ان $f(x) = 6x(1-x) , 0 < x < 1$ تمثل دالة الكثافة الاحتمالية الى X . جد المنوال لهذا التوزيع .

الحل :

$$f'(x) = 6(1-2x)$$

وبجعل $f'(x) = 0$ نحصل على :

$$6(1-2x) = 0 \rightarrow 1 - 2x = 0 \rightarrow x = \frac{1}{2}$$

كذلك فان $f''(x) = -12$ فاذن $f''(x)\Big]_{x=\frac{1}{2}} = -12 < 0$

من ذلك نستنتج ان المنوال في هذا التوزيع هو $x = \frac{1}{2}$

لاحظ من هذا المثال ان $x = \frac{1}{2}$ قيمة معرفة في الفترة $(0 , 1)$. ان القيمة العظمى للدالة $f(x)$ هي .

$$\text{Max } f(x) = f(x)\Big]_{x=\frac{1}{2}} = \frac{3}{2}$$

مثال (٢) : افرض ان $f(x) = \frac{1}{2} x^2 e^{-x} ; x > 0$ تمثل دالة الكثافة الاحتمالية للمتغير X . جد المنوال في هذا التوزيع .

الحل :

$$f'(x) = \frac{1}{2} e^{-x}(-x^2 + 2x)$$

وبجعل $f'(x) = 0$ نحصل على

$$\frac{1}{2} e^{-x}(-x^2 + 2x) = 0 \rightarrow e^{-x} . x(x - 2) = 0$$

واضح ان هنالك ثلاثة حلول لهذه المعادلة هي :

$e^{-x} = 0$ فذلك يعني ان $x \rightarrow \infty$ او $x = 0$ او ان $x - 2 = 0$ فذلك يعني ان $x = 2$. نجد المشتقة الثانية وهي $f''(x) = \frac{1}{2} e^{-x}(x^2 - 4x + 2)$ ثم نختبر $f''(x)$ عند قيم x التي حصلنا عليها من حل $f'(x) = 0$ وعلى النحو التالي :

لقيمة $x \rightarrow \infty$ فان $\lim_{x \rightarrow \infty} f''(x) = 0$ اي ان $x \rightarrow \infty$ حل غير معقول . ولقيمة $x = 0$ فان $f''(x) = 1$ وهذا ايضاً هو حل غير معقول بسبب ان $f''(0) > 0$ ولقيمة $x = 2$ فان $f''(x) = -e^{-2} < 0$ وهذا حل معقول بسبب ان $f''(x) < 0$ وان $x = 2$ قيمة معرفة في الفترة $[0, \infty)$. عليه فان قيمة المنوال في هذا التوزيع هي $x = 2$ وبذلك فان القيمة العظمى للدالة $f(x)$ هي :

$$\text{Max } f(x) = f(x)]_{x=2} = 2e^{-2}$$

مثال (٣) : افرض أن x متغير عشوائي بدالة كتلة احتمالية $p(x) = \frac{2^x e^{-2}}{x!} \quad ; x = 0, 1, \ldots$ جدول المنوال في هذا التوزيع .

الحل : نجد التوزيع الاحتمالي للمتغير x أي :

x :	0	1	2	3	4	5	6	...
p(x) :	e^{-2}	$2e^{-2}$	$2e^{-2}$	$\frac{4}{3}e^{-2}$	$\frac{2}{3}e^{-2}$	$\frac{4}{15}e^{-2}$	$\frac{4}{45}e^{-2}$	...

نستنتج ان

$$P(0) < P(1) = P(2) > P(3) > P(4) > \ldots$$

وهذا يعني ان هذا التوزيع يمتلك منوالين هما $x = 1$, $x = 2$.

مثال (٤) : افرض ان $x = 1, 2, ..., 6$; $p(x) = \frac{1}{6}$ تمثل دالة الكتلة الاحتمالية للمتغير X . جد المنوال لهذا التوزيع

الحل : حيث ان $p(1) = p(2) = ... = p(6) = \frac{1}{6}$

فذلك يعني ان المنوال في هذا التوزيع غير موجود بسبب ان التوزيع منتظم في تخصيص الكتل الاحتمالية لعناصر x .

٣ ـ ٢ : الوسيط Median .

يعد الوسيط هو الاخر احد مقاييس النزعة المركزية ذات قيمة معرفة على المحور السيني . ويعرف الوسيط بأنه قيمة من قيم المتغير العشوائي X التي تقسم المساحة تحت منحنى الدالة $f(x)$ الى قسمين متساويين . وحيث ان $f(x)$ هي دالة كثافة احتمالية فذلك يعني ان الوسيط يمثل تلك القيمة التي تجعل نصف الاحتمال المقترن بفضاء العينة Ω (البالغ واحد) واقع الى يمينها والنصف الاخر الى يسارها . وبفرض ان M تمثل الوسيط وان Ω تمثل مجموعة الاعداد الحقيقية . فان M تمثل قيمة من قيم X التي تحقق

المعادلة التكاملية التالية : $\int_{-\infty}^{M} f(x)\,dx = \frac{1}{2}$ وحيث ان

$\int_{-\infty}^{M} f(x)\,dx = P_r(X \leq M) = F(M)$ فذلك يعني انه يمكن الحصول على قيمة

الوسيط من خلال الدالة التوزيعية $F(x)$ كنتيجة لحل الصيغة $F(M) = \frac{1}{2}$ نسبة الى M . مما تقدم نلاحظ ان الوسيط يتمثل بقيمة واحدة فقط بعكس المنوال . وان وحدات قياس الوسيط هي نفس وحدات قياس المتغير X . اما في حالة المتغيرات المتقطعة فان الوسيط يمثل تلك القيمة (قد تكون معرفة في Ω او قد لاتكون) التي تقسم الاحتمال الكلي المقترن بفضاء العينة الى قسمين متساويين نصفه الى يمين الوسيط والنصف الاخر الى يساره . وهذا يعني ان قيمة الوسيط M يمكن الحصول عليها من حل الصيغة $F(M) = \frac{1}{2}$ نسبة الى M . وعموماً فان الهدف من دراسة الوسيط هو تكوين فكرة عن القيمة التي تشطر الاحتمال المقترن بفضاء العينة للمتغير

العشوائي الى قسمين متساويين اضافة الى كونه مقياس بديل للمتوسط في حالة عدم امكانية ايجاد الاخير . وسوف نلاحظ في الفصل التاسع اسلوب اشتقاق التوزيع الاحتمالي الى الوسيط .

مثال (٥) : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = e^{-x}, x \geq 0$ جد الوسيط لهذا التوزيع .

الحل :

$$\int_0^M e^{-x}\,dx = (1 - e^{-M}) = \frac{1}{2}$$

وهذا يعني ان $e^{-M} = \frac{1}{2}$ فاذن $M = -\ln\left(\frac{1}{2}\right) \simeq 0{\cdot}6931471$

مثال (٦): لتكن $f(x)=2x; 0<x<1$ تمثل دالة الكثافة الاحتمالية للمتغير X .
جد الوسيط في هذا التوزيع .

الحل :

$$\int_0^M 2x\,dx = M^2 = \frac{1}{2} \qquad \therefore M = \pm\frac{1}{\sqrt{2}}$$

وحيث ان $M=-\frac{1}{\sqrt{2}}$ قيمة غير معرفة في الفترة [0 , 1] نستنتج ان الوسيط في هذا التوزيع هو $M = \frac{1}{\sqrt{2}}$.

مثال (٧) :لتكن $p(x)=ab^x ; x=0, 1, 2, \ldots$, $a+b=1$, تمثل دالة الكتلة الاحتمالية للمتغير X . جد الوسيط في هذا التوزيع .

الحل :

$$\sum_{x=0}^{M} ab^x = a \sum_{x=0}^{M} b^x = a (1 + b + b^2 + ... + b^M)$$

لكن المجموع اعلاه يمثل مجموع حدود متوالية هندسية اساسها مساوٍ الى b فاذن $\sum_{x=0}^{M} b^x = \frac{1 - b^{M+1}}{1 - b}$ وهذا يعني ان

$$a \sum_{x=0}^{M} b^x = 1 - b^{M+1} \quad , \quad a = 1 - b .$$

فاذن :

$$1 - b^{M+1} = \frac{1}{2} \qquad \therefore b^{M+1} = \frac{1}{2}$$

او ان

$$(M + 1) \log b = \log\left(\frac{1}{2} \right), \quad \therefore M + 1 = \frac{\log\left(\frac{1}{2} \right)}{\mathrm{Log} (b)}$$

فاذن :

$$M = \frac{\mathrm{Log}\left(\frac{1}{2} \right)}{\mathrm{Log} (b)} - 1 = C - 1$$

ويلاحظ ان $C < 1$ عندما $b < \frac{1}{2}$ وهذا يعني ان $M < 0$ وهذا غير جائز . في حين يلاحظ ان $C \geq 1$ عندما $b \geq \frac{1}{2}$ وهذا يعني ان $M \geq 0$. نستنتج ان قيمة M تتحدد عندما $b \geq \frac{1}{2}$.

٣ ـ ٣ : الرُبيعات Quartiles

في حالة تجزئة الاحتمال الكلي المقترن بفضاء العينة الى اربعة اجزاء متساوية فان قيم المتغير العشوائي X الثلاث التي تجزأ هذا الاحتمال تسمى الربيعات . فاذا كان X متغير عشوائي مستمر معرفة قيمه في مجموعة الاعداد الحقيقية R وان Q_1، Q_2 ، Q_3 تمثل الربيعات الثلاث عندئذٍ يمكن تحديد قيمة كل من Q_1, Q_2, Q_3 وفق مايلي:

$$\int_{-\infty}^{Q_1} f(x)\,dx = \frac{1}{4}, \int_{-\infty}^{Q_2} f(x)\,dx = \frac{1}{2}, \int_{-\infty}^{Q_3} f(x)\,dx = \frac{3}{4}$$

واضح مما تقدم ان Q_2 ماهو الا الوسيط في التوزيع . واذا كانت الدالة التوزيعية $F(x)$ معلومة فانه يمكن حساب قيم الربيعات وفق مايلي :

$$F(Q_1) = \frac{1}{4}, F(Q_2) = \frac{1}{2}, F(Q_3) = \frac{3}{4}$$

وبشكل عام فان قيمة الربيع i تمثل قيمة من قيم X التي تحقق المعادلة التكاملية :

$$\int_{-\infty}^{Q_i} f(x)\,dx = F(Q_i) = \frac{i}{4}, i = 1, 2, 3 .$$

ونفس المفهوم اعلاه ينطبق على حالة المتغيرات العشوائية المتقطعة بمجرد الاستعاضة عن رمز التكامل برمز الجمع .

مثال (٨) : لتكن $f(x)=\frac{1}{x^2}; x \geq 1$ تمثل دالة الكثافة الاحتمالية للمتغير العشوائي X . جد الربيعات في هذا التوزيع .

الحل : للسهولة في انجاز الحل نجد اولاً الدالة التوزيعية .

$$F(x) = \int_{1}^{x} \frac{1}{u^2}\,du = 1 - \frac{1}{x}$$

فاذن

$$F(Q_1) = 1 - \frac{1}{Q_1} = \frac{1}{4} \rightarrow \frac{1}{Q_1} = \frac{3}{4} \quad \therefore Q_1 = \frac{4}{3}$$

$$F(Q_2) = 1 - \frac{1}{Q_2} = \frac{1}{2} \rightarrow \frac{1}{Q_2} = \frac{1}{2} \quad \therefore Q_2 = 2$$

$$F(Q_3) = 1 - \frac{1}{Q_3} = \frac{3}{4} \rightarrow \frac{1}{Q_3} = \frac{1}{4} \quad \therefore Q_3 = 4$$

مثال (٩) : اذا علمت ان الدالة التوزيعية لمتغير عشوائي X هي $F(x) = 1 - e^{-x}, x \geq 0$ جد الربيعات في هذا التوزيع .

الحل :

$$F(Q_1) = 1 - e^{-Q_1} = \frac{1}{4} \rightarrow e^{-Q_1} = \frac{3}{4} \quad \therefore Q_1 = -\ln\left(\frac{3}{4}\right)$$

$$F(Q_2) = 1 - e^{-Q_2} = \frac{1}{2} \rightarrow e^{-Q_2} = \frac{1}{2} \quad \therefore Q_2 = -\ln\left(\frac{1}{2}\right)$$

$$F(Q_3) = 1 - e^{-Q_3} = \frac{3}{4} \rightarrow e^{-Q_3} = \frac{1}{4} \quad \therefore Q_3 = -\ln\left(\frac{1}{4}\right)$$

٣ ـ ٤ : العُشيرات Deciles

لايختلف مفهوم العشيرات عن مفهوم الربيعات سوى انه في هذه الحالة يتم تجزئة الاحتمال الكلي المقترن بفضاء العينة Ω الى عشرة اجزاء متساوية وعندئذٍ فان قيم المتغير العشوائي X التسع التي تجزأ هذا الاحتمال تسمى العشيرات . فاذا كان X متغير عشوائي مستمر معرفة قيمه في مجموعة الاعداد الحقيقية R وان $i = 1, 2, ..., 9$

، D_i يمثل العُشير i عندئذٍ يمكن الحصول على قيمة D_i من خلال حل المعادلة التكاملية التالية :

$$\int_{-\infty}^{D_i} f(x)\,dx = F(D_i) = \frac{i}{10}\ , i = 1, 2, \ldots, 9.$$

مع ملاحظة ان D_5 ماهي الا قيمة الوسيط في التوزيع .

مثال : لمعطيات المثال (٨) الوارد في الفقرة (٣ ـ ٣) جد صيغة لحساب قيم العشيرات في التوزيع .

الحل : ان $F(x) = 1 - \frac{1}{x}$ فاذن ،

$$F(D_i) = 1 - \frac{1}{D_i} = \frac{i}{10}\ , i = 1, 2, \ldots, 9$$

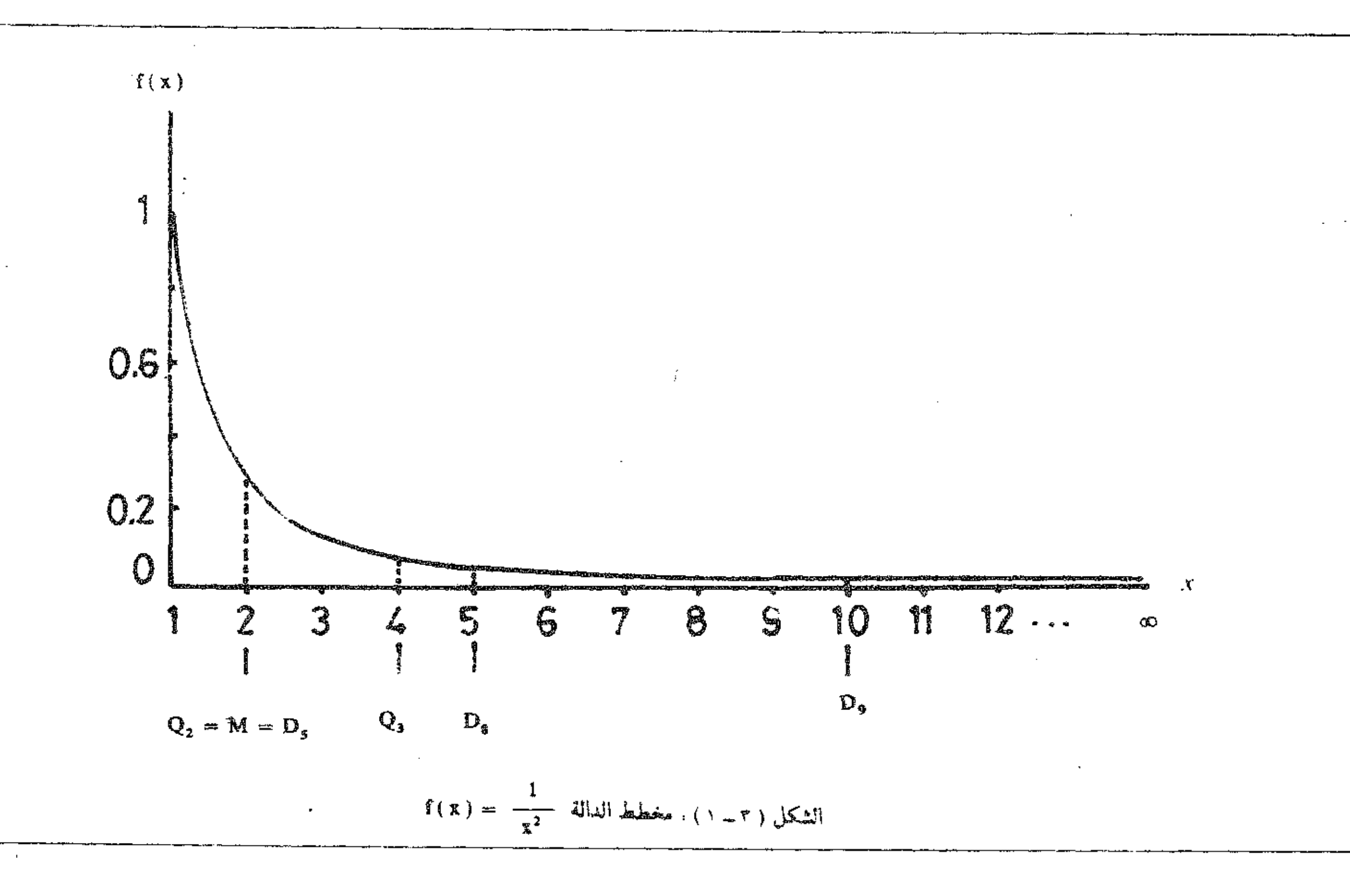
عليه فان

$$\frac{1}{D_i} = 1 - \frac{i}{10} = \frac{10 - i}{10}$$

فاذن

$$D_i = \frac{10}{10 - i}\ , i = 1, 2, \ldots, 9.$$

والشكل (٣ ـ ١) يوضح مخطط الدالة $f(x) = \frac{1}{x^2}$ وموقع بعض الربيعات والعشيرات .



الشكل (٣ ـ ١) : مخطط الدالة $f(x) = \frac{1}{x^2}$

مثال (١٠) : لمعطيات المثال (٩) الوارد في الفقرة (٣ – ٣) جد صيغة لحساب قيم العشيرات في ذلك التوزيع .

الحل : ان $F(x) = 1 - e^{-x}\ ,\ x \geq 0$. فاذن

$$F(D_i) = 1 - e^{-D_i} = \frac{i}{10}\ ,\ i = 1, 2, \ldots, 9.$$

فاذن

$$e^{-D_i} = 1 - \frac{i}{10} = \frac{10 - i}{10}$$

اي ان

$$D_i = -\ln\left(\frac{10 - i}{10}\right), i = 1, 2, \ldots, 9.$$

٣ ـ ٥ : الانحراف الربيعي Quartile deviation

يعد الانحراف الربيعي احد مقاييس التشتت المطلقة . الهدف منه قياس درجة تشتت قيم المتغير العشوائي في توزيع احتمالي معين . ويعرف الانحراف الربيعي بانه متوسط الفرق مابين الربيع الثالث والربيع الاول . فاذا رمزنا للانحراف الربيعي بالرمز Q فان : $Q = \frac{Q_3 - Q_1}{2}$.

وعلى الرغم من ان هذا المقياس غير دقيق في قياس درجة التشتت (كالتباين والانحراف المطلق) الا انه مفيد في تلك الحالات التي يتعذر فيها حساب التباين او الانحراف المطلق . كذلك يلاحظ ان وحدات قياس الانحراف الربيعي هي نفس وحدات قياس المتغير العشوائي X .

مثال (١١) : لمعطيات المثال (٨) الوارد في الفقرة (٣ ـ ٣) جد الانحراف الربيعي .

الحل : ان $Q_3 = 4 \cdot Q_1 = \frac{4}{3}$. فاذن $Q = \frac{4 - \frac{4}{3}}{2} = \frac{4}{3}$

٣ ـ ٦ : معامل الاختلاف Coefficient of variation (C.V)

يعتبر معامل الاختلاف احد مقاييس التشتت النسبية . وهو مقياس مفيد في حالة اجراء المقارنة بين توزيعين مختلفين من حيث الوسط والتباين بهدف معرفة اي منهما قيمه اكثر تجانساً . ويعرف معامل الاختلاف بأنه النسبة بين الانحراف المعياري في توزيع معين الى وسط ذلك التوزيع . اي ان

$$C.V = \frac{\sigma_x}{\mu_x} . 100$$

ويلاحظ من هذه الصيغة ان معامل الاختلاف مقياس خال من وحدات القياس . وفي حالة تعذر حساب عزوم توزيع معين عندئذٍ يستعاض عن معامل الاختلاف بمعامل آخر يسمى معامل التشتت Coefficient of dispersion الذي يعتمد على قيمة الربيع الاول والربيع الثالث . وصيغة هذا المعامل هي :

$$C.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} . 100$$

مثال (١٢) : اذا علمت ان الوسط لتوزيع معين كان 10 والانحراف المعياري كان 5 في حين ان الوسط لتوزيع آخر كان 15 والانحراف المعياري كان 6 . اي من هذين التوزيعين قيمة اكثر تجانساً .

الحل : لاول وهلة وعلى اساس الانحراف المعياري نلاحظ ان قيم التوزيع الاول اكثر تجانساً من قيم التوزيع الثاني . على اساس معامل الاختلاف نلاحظ مايلي :

$$C.V_1 = \frac{5}{10} . 100 = 50\% , C.V_2 = \frac{6}{15} . 100 = 40\%$$

وحيث ان معامل الاختلاف في التوزيع الثاني اقل من معامل الاختلاف في التوزيع الاول فذلك يعني ان قيم التوزيع الثاني اكثر تجانساً من قيم التوزيع الاول .

مثال (١٣) : لمعطيات المثال (٨) الوارد في الفقرة (٣ ـ ٣) جد معامل التشتت .

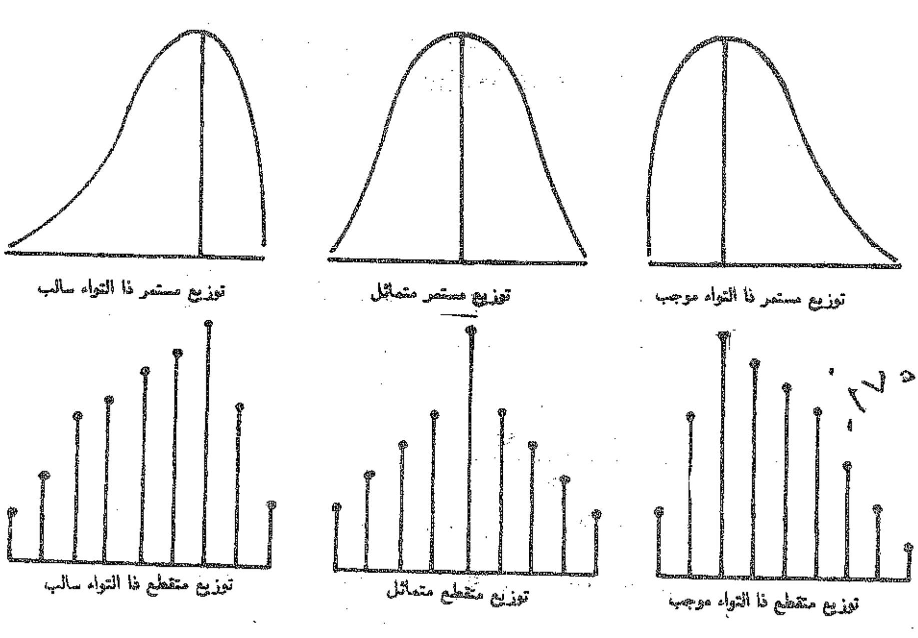
الحل : حيث ان $Q_3 = 4 , Q_1 = \frac{4}{3}$ فاذن

$$C.D = \frac{4 - 4/3}{4 + 4/3} . 100 = 50\%$$

٣ـ ٧ : الالتواء Skewness

تنقسم التوزيعات الاحتمالية بشكل عام الى قسمين رئيسين الاول منها يدعى التوزيعات المتماثلة Symmetric distributions والثاني يدعى التوزيعات الملتوية Skewed distribution . ويقصد بالتوزيعات المتماثلة بانها تلك التوزيعات التي تكون المساحة تحت منحنى دالة التوزيع الى يمين المنوال مساوية ومشابهة لتلك الى يساره (في حالة التوزيعات المستمرة) . اما في حالة التوزيعات المتقطعة فان حالة التماثل في التوزيع تتحقق عندما تكون الكتل الاحتمالية المقترنة بعناصر الفضاء Ω المتناظرة من حيث الموقع حول قيمة المنوال (يميناً ويساراً) متساوية القيمة . وفي غير هذه الاحوال يقال ان التوزيع ملتوٍ . والالتواء على نوعين : التواء موجب والتواء سالب . ويقال ان التوزيع ملتو التواء موجب اذا كانت المساحة (او مجموع الكتل الاحتمالية) الى يمين المنوال اكبر من تلك الى يساره . في حين يقال ان التوزيع ملتو التواء سالب اذا كانت المساحة (او مجموع الكتل الاحتمالية) الى يمين المنوال اقل من تلك الى يساره .

والشكل (٣ ـ ٢) يوضح اشكال مختلفة لتوزيعات احتمالية :



الشكل (٣ ـ ٢) : اشكال مختلفة لتوزيعات احتمالية .

ومن وجهة النظر الرياضية وبفرض ان X متغير عشوائي مستمر بدالة كثافة احتمالية f(x) معرفة قيمة في مجموعة الاعداد الحقيقية R . وبفرض ان a عدد موجب عندئذٍ يقال ان التوزيع الاحتمالي متماثل اذا وفقط اذا كانت $f(-a)=f(a)$ وفي غير ذلك يقال ان التوزيع ملتوٍ. اما في حالة المتغيرات العشوائية المتقطعة وبفرض ان $X = m$ يمثل المنوال الوحيد في التوزيع ، عندئذٍ يقال ان التوزيع الاحتمالي متماثل اذا وفقط اذا كانت $P(X = m-i)=P(X=m+i), i = 1, 2, \ldots$. اما اذا كان التوزيع الاحتمالي المتقطع يمتلك منوالين هما $X = m - 1$ ، $X = m$ عندئذٍ يقال ان التوزيع الاحتمالي متماثل اذا وفقط اذا كان $P(X = m - 1 - i) = P(X = m + i), i = 1, 2, \ldots$. واذا كان التوزيع الاحتمالي متماثلاً عندئذٍ فان مقاييس النزعة المركزية الثلاث (الوسط، الوسيط، المنوال) تكون متساوية القيمة. في حين اذا كان التوزيع الاحتمالي ذا التواء موجب فان الوسط > الوسيط > المنوال . والعكس صحيح اي اذا كان التوزيع الاحتمالي ذا التواء سالب فان الوسط < الوسيط < المنوال . ويمكن قياس درجة ونوع الالتواء لتوزيع احتمالي وفق المعاملات التالية :

١ ـ معامل التواء التوزيع لكارل بيرسون

Karl pearson's coefficient of skewness

ان صيغة هذا المعامل هي :

$$S_k = \frac{\mu_x - M_0}{\sigma_x}$$

حيث M_0 تعني المنوال لذلك التوزيع الاحتمالي الذي وسطه هو μ_x وتباينه σ_x^2 . وفي حالة تعذر حساب قيمة M_0 يمكن استخدام الصيغة التالية لحساب معامل الالتواء وهي :

$$S_k = \frac{3(\mu_x - M_e)}{\sigma_x}$$

حيث M_e تعني الوسيط للتوزيع الاحتمالي . وهنالك معامل آخر يستند الى العزوم . صيغة هذا المعامل هي $S_k = \frac{\mu_3^2}{\sqrt{\mu_2^3}}$ حيث μ_2 ، μ_3 هما على التوالي العزم المركزي الثاني والثالث للتوزيع الاحتمالي .

٢ ـ معامل التواء التوزيع لبوْلي
Bowley's coefficient of skewness

حالة تعذر حساب عزوم التوزيع بسبب عدم تحقق خاصية التقارب المطلق، عندئذٍ يمكن استخدام المعامل التالي الذي تستند صيغته على قيم الربيعات وهي :

$$S_k = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

واياً كان معامل الالتواء المستخدم اذا لوحظ ان :

$S_k < 0$: فذلك يعني ان التوزيع الاحتمالي ذو التواء سالب .

$S_k = 0$: فذلك يعني ان التوزيع الاحتمالي متماثل .

$S_k > 0$: فذلك يعني ان التوزيع الاحتمالي ذا التواء موجب .

وتزداد شدة التواء التوزيع كلما ابتعدت $|S_k|$ عن الصفر .

مثال (١٤) : اذا علمت ان قيمة الوسط في توزيع احتمالي كانت 4 والمنوال كان 5.2 . والانحراف المعياري كان 3 . حدد درجة ونوع التواء هذا التوزيع .

الحل :

$$S_k = \frac{4 - 5{\cdot}2}{3} = -0{\cdot}4$$

وهذا يعني ان التوزيع ذو التواء سالب وان درجة الالتواء هي 0.4 .

مثال (١٥) : لمعطيات المثال (٩) الوارد في الفقرة (٣ ـ ٣) حدد درجة ونوع التواء هذا التوزيع .

الحل : ان

$$Q_1 = 0{\cdot}2877 ,\ Q_2 = 0{\cdot}6931 ,\ Q_3 = 1{\cdot}3863$$

فاذن

$$S_k = \frac{1{\cdot}3863 + 0{\cdot}2877 - 2(0{\cdot}6931)}{1{\cdot}3863 - 0{\cdot}2877} = 0{\cdot}262$$

وهذا يعني ان التوزيع ذا التواء موجب وان درجة التوائه هي 0.262 .

مثال (١٦) اذا كان $S_k = \frac{\mu_x - M_0}{\sigma_x} = \frac{3(\mu_x - M_e)}{\sigma_x}$. حيث ان μ_x ، M_0 ، M_e هي على التوالي الوسط والمنوال والوسيط في توزيع احتمالي لمتغير عشوائي مستمر . وبفرض ان $S_k > 0$. برهن ان $\mu_x > M_e > M_0$.

البرهان : حيث ان $S_k > 0$ فذلك يعني ان ،

$$\frac{\mu_x - M_0}{\sigma_x} > 0 \rightarrow \mu_x > M_0 \quad \dots (1)$$

وان

$$\frac{3(\mu_x - M_e)}{\sigma_x} > 0 \rightarrow \mu_x > M_e \quad \dots (2)$$

كذلك فان :

$$\mu_x = M_0 = 3\mu_x - 3M_e$$

اي ان

$$2\mu_2 = 3M_e - M_0 \quad \dots (3)$$

من العلاقة (2) نلاحظ ان

$$2\mu_x > 2M_e \quad \dots (4)$$

وبالتعويض عن (4) في (3) نجد ان

$$2M_e < 3M_e - M_0 \rightarrow M_0 < M_e$$

عليه نستنتج ان $\mu_x > M_e > M_0$.

٣ ـ ٨ : التفلطح Kurtosis

يعرف التفلطح بأنه مقدار تسطح flatness او تدبب Peakedness منحنى التوزيع الاحتمالي لمتغير عشوائي . ويرتبط مفهوم التفلطح ارتباطاً وثيقاً مع مفهوم التشتت . فكلما كان تشتت قيم المتغير عالياً فذلك مؤشر لتسطح منحنى التوزيع

الاحتمالي . ويمكن قياس درجة تفلطح منحنى التوزيع الاحتمالي وفق الصيغة التالية المقترحة من قبل العالم كارل بيرسون وهي :

$$\beta = \frac{\mu_4}{\mu_2^2} \quad \text{حيث} \quad \lambda = \beta - 3$$

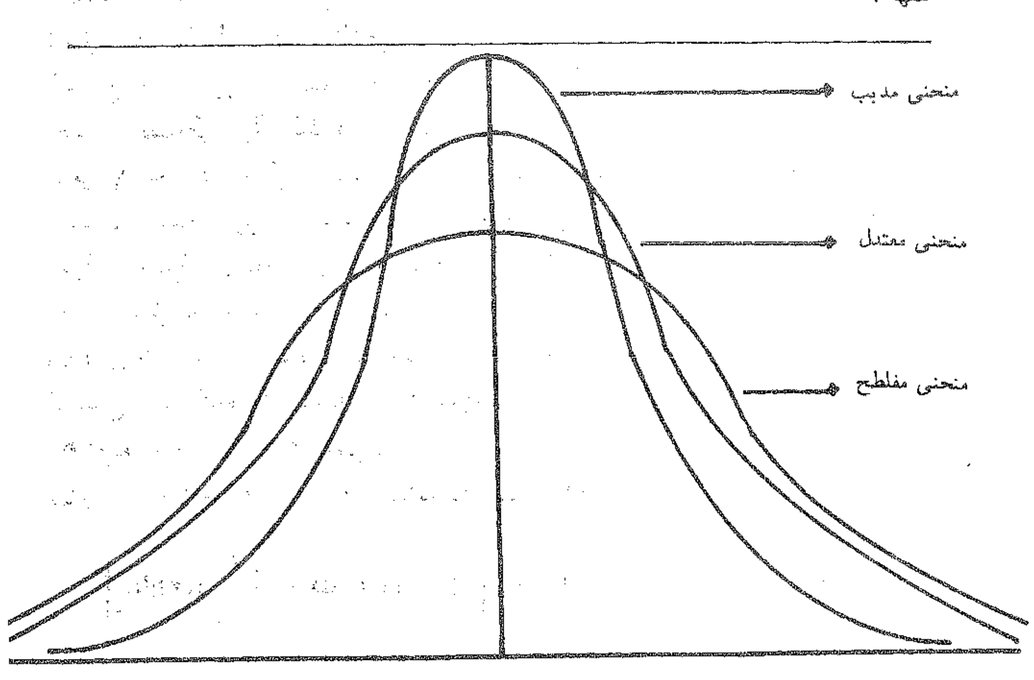
μ_2 و μ_4 هما على التوالي العزم المركزي الثاني والرابع . فاذا كانت :

$\lambda > 0$: فذلك يعني ان منحنى التوزيع الاحتمالي مدبب وتزداد درجة التدبب بزيادة قيمة λ . واذا كانت

$\lambda = 0$: فذلك يعني ان منحنى التوزيع الاحتمالي معتدل التفلطح . واذا كانت

$\lambda < 0$: فذلك يعني ان منحنى التوزيع الاحتمالي مفلطح وتزداد درجة التفلطح بانخفاض قيمة λ .

والشكل (٣ ـ ٣) يوضح مقارنة بين ثلاثة منحنيات حسب درجة تفلطح كل منها :



الشكل (٣ ـ ٣) ، مقارنة بين ثلاثة منحنيات حسب درجة التفلطح .

مثال (١٧) : لوحظ في توزيع احتمالي ان قيمة العزم المركزي الثاني كانت 2 وقيمة العزم المركزي الرابع كانت 14 . ما هي درجة ونوع تفلطح منحني هذا التوزيع .

الحل : $\mu_4 = 14 \cdot \mu_2 = 2$ فاذن $\beta = \frac{14}{4} = 3.5 , \lambda = 0.5$ وحيث ان $\lambda > 0$ فذلك يعني ان منحنى هذا التوزيع مدبب وان درجة تفلطحه هي 0.5 .

٣ ـ ٩ : التوزيعات المقطوعة (المبتورة)

truncated distributions

يتطلب الامر في بعض الاحيان استنتاج دالة كثافة احتمالية (او دالة كتلة احتمالية) لمتغير عشوائي X معرفة قيمه على جزء من القيم المعرفة في Ω لاسباب تتعلق بطبيعة الدراسة او التجربة . وهذا يعني ان هنالك قطعاً (او بتراً) في التوزيع الاحتمالي . ان عملية البتر في التوزيعات الاحتمالية تؤثر بشكل مباشر على احدى خصائص دوال الكثافة او الكتلة الاحتمالية وهي ان الاحتمال المقترن بفضاء المتغير X بعد عملية البتر سوف يكون اقل من واحد ، اي $P_X(\Omega) < 1$ الامر الذي يتطلب اشتقاق توزيع جديد من التوزيع الاصلي يحقق خصائص هذا النوع من الدوال . وسوف نوضح اسلوب الاشتقاق لحالة المتغيرات العشوائية المستمرة والاسلوب ذاته ينطبق لحالة المتغيرات المتقطعة . ليكن X متغير عشوائي بدالة كثافة احتمالية $f(x)$ وان قيم X معرفة بمجموعة الاعداد الحقيقية R . وافرض اننا نرغب في استنتاج دالة كثافة احتمالية الى X معرفة قيمه على مجموعة جزئية من Ω لتكن $\Omega^* = \{x : a < x < b\}$ حيث a , b عددان معرفان في Ω . لتكن $F(x)$ تمثل الدالة التوزيعية الى X مشتقة على اساس $f(x)$. ليكن $c > 0$. عندئذٍ

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx = c\left[\int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)dx\right]$$

$$= c[F(b) - F(a)]$$

وحتى تكون الدالة الجديدة تتصف بخصائص دوال الكثافة الاحتمالية فذلك

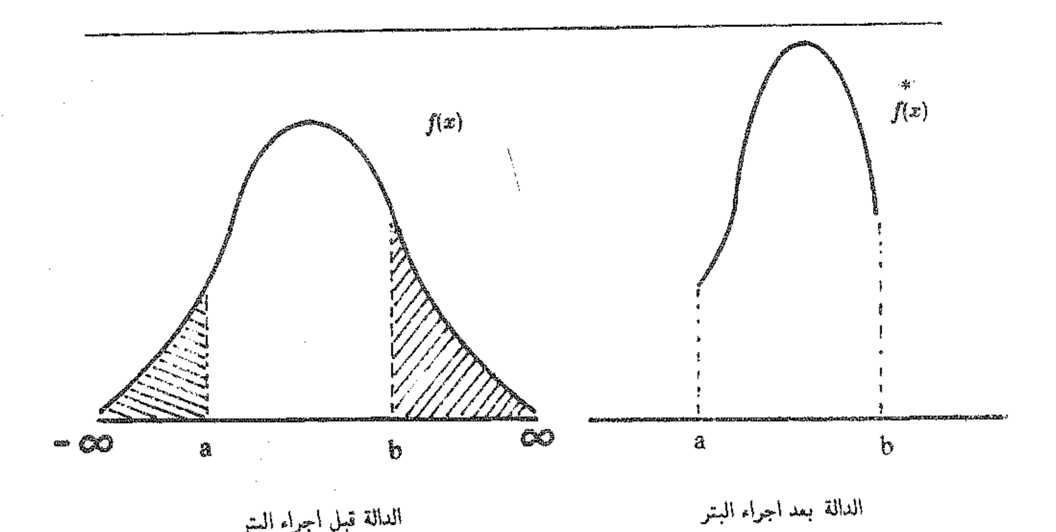
يعني ان $c\int_a^b f(x)\,dx = 1$. عليه فان $c[F(b) - F(a)] = 1$ وان

$c = [F(b) - F(a)]^{-1}$ عليه نستنتج ان دالة الكثافة الاحتمالية الى X

بعد اجراء عملية البتر (اي اهمال قيم الفترتين $(-\infty, a]$, $[b, \infty)$)

هي : $f^*(x) = [F(b) - F(a)]^{-1} . f(x) \quad ; \quad a < x < b$

وكما هو موضح في الشكل (٣ ـ ٤) :



الشكل (٣ ـ ٤) ، توضيح لعملية البتر في التوزيعات الاحتمالية .

وهذا يعني ان دالة الكثافة الاحتمالية الجديدة للمتغير X ماهي الا دالة الكثافة الاحتمالية الاصلية مقسومة على احتمال الفترة $[a, b]$.وحيث ان $P_r(a < X < b) < 1$ فذلك يعني ان $1 < [F(b) - F(a)]^{-1}$ عليه وبشكل عام

$$\mathrm{Max}\, f^*(x) > \mathrm{Max}\, f(x)$$

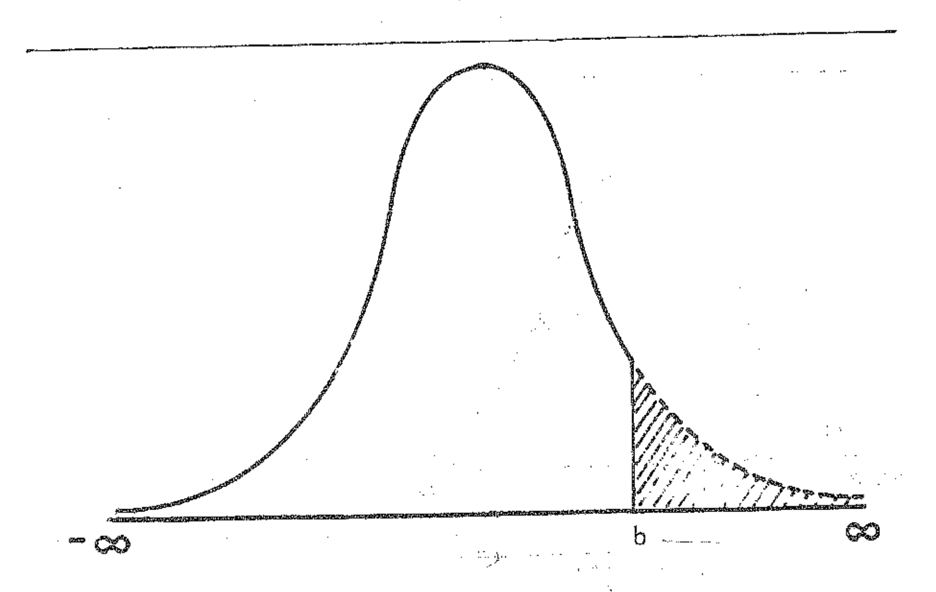
وهنالك حالتان خاصتان للبتر يمكن استنتاجهما من الحالة العامة اعلاه . وهما :

١ ـ البتر من الجانب الايمن :

في هذه الحالة نرغب في استنتاج دالة الكثافة الاحتمالية للمتغير X معرفة قيمة على الفترة [b , ∞ –) . وهذا يعني ان :

$$f^*(x) = [F(b) - F(-\infty)]^{-1} . f(x) = \frac{f(x)}{F(b)} ; \quad -\infty < x < b$$

وكما هو موضح في الشكل (٣ ـ ٥) .



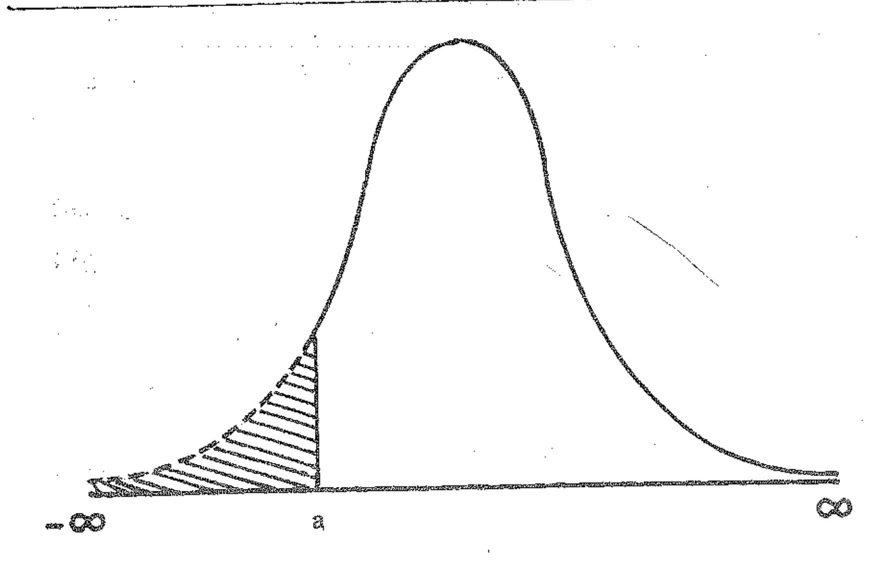
الشكل (٣ ـ ٥) : البتر في الدالة f (x) من الجانب الايمن .

٢ – البتر من الجانب الايسر .

في هذه الحالة نرغب في استنتاج دالة الكثافة الاحتمالية للمتغير X معرفة قيمة على الفترة [a , ∞] وهذا يعني ان :

$$f^*(x) = [F(\infty) - F(a)]^{-1} . f(x) = [1 - F(a)]^{-1} . f(x) ; x > a$$

وكما هو موضح في الشكل (٣ ـ ٦) :



الشكل (٣ـ ٩) ، البتر في الدالة $f(x)$ من الجانب الايسر

ملاحظة : بعد ان يتم استنتاج دالة الكثافة الاحتمالية (او دالة الكتلة الاحتمالية) بعد اجراء عملية البتر في التوزيع ، يمكن وعلى ضوء تلك الدالة ايجاد كل مايتعلق بالتوزيع الجديد من مقاييس وعزوم ودوال توليد العزوم وغيرها دون اللجوء الى التوزيع الاصلي .

مثال (١٨) : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = e^{-x}$; $x \geq 0$

جد دالة الكثافة الاحتمالية للمتغير X معرفة قيمه على الفترة $[2, \infty]$ ثم جد :
أ ــ الدالة التوزيعية بعد البتر ، ب ــ الدالة المولدة للعزوم حول نقطة الاصل بعد البتر ، جـ ــ الوسط والتباين بعد البتر .

الحل : نجد اولاً الدالة التوزيعية قبل البتر وهي :

$$F(x) = \int_0^x e^{-u}\,du = 1 - e^{-x} \quad ; x \geq 0$$

ان المطلوب هو اجراء البتر من الجانب الايسر . فاذن

$$f^*(x) = [1 - F(2)]^{-1} . e^{-x} = [e^{-2}]^{-1} . e^{-x} = e^2 . e^{-x} = e^{2-x} ; x \geq 2$$

لاحظ ان

$$\int_2^\infty f^*(x)\,dx = \int_2^\infty e^{2-x}\,dx = 1$$

فاذن

أ ـ

$$F^*(x) = \int_2^x e^{2-u}\,du = -[e^{2-u}]_2^X = 1 - e^{2-x} ; x \geq 2$$

لاحظ ان

$$F^*(\infty) = 1 , F^*(2) = 0$$

ب ـ

$$M_X^*(t) \int_2^\infty e^{tx} . e^{2-x}\,dx = e^2 \int_2^\infty e^{-x(1-t)}\,dx$$

$$= e^2 . \frac{e^{-2(1-t)}}{(1-t)} = \frac{e^{2t}}{(1-t)} ; t < 1$$

لاحظ ان $M_X^*(0) = 1$.

جـ ـ لغرض ايجاد الوسط والتباين نجد اولاً الدالة المولدة التراكمية . اي

$$K_X^*(t) = \ln M_X^*(t) = 2t - \ln(1-t)$$

فاذن

$$K_X^{*\prime}(t) = 2 + \frac{1}{1-t}$$

وهذا يعني ان

$$EX = K_X^{*\prime}(0) = 3$$

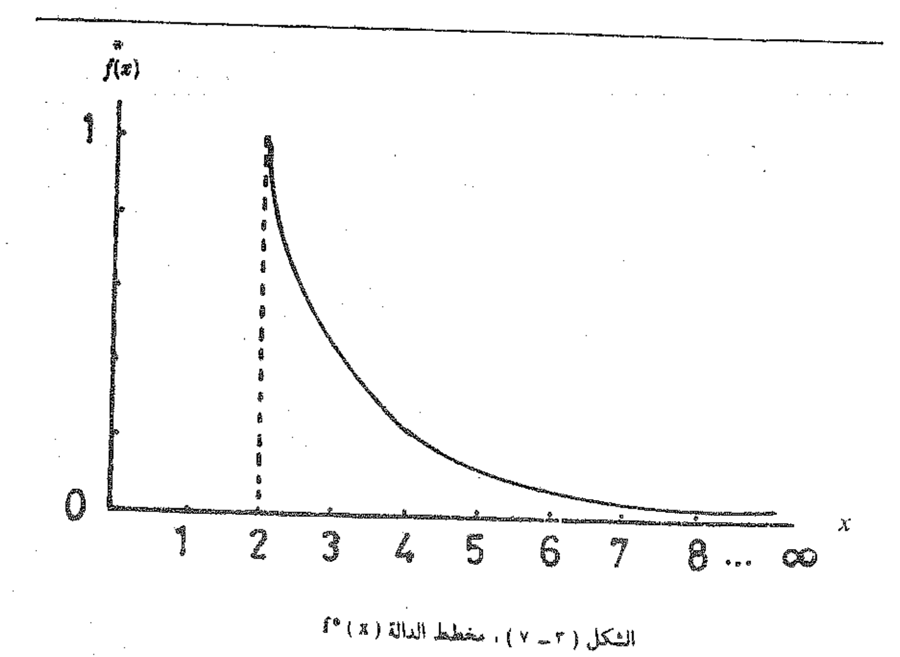
وان

$$K_X^{*\prime\prime}(t) = \frac{1}{(1-t)^2}$$

فاذن

$$V(X) = K_X^{*\prime\prime}(0) = 1$$

والشكل (٣ ـ ٧) يوضح مخطط الدالة $f^*(x)$.



الشكل (٣ ـ ٧) ، مخطط الدالة $f^*(x)$

مثال (١٩) : افرض ان X متغير عشوائي بدالة كتلة احتمالية $x = 0, 1, 2, \ldots$ ، $P(x) = \frac{2^x e^{-2}}{x!}$; جد دالة الكتلة الاحتمالية الى X معرفة قيمه بالمجموعة $\{x : x = 3, 4, 5, \ldots\}$.

الحل : ان

$$F(2) = P(0) + P(1) + P(2) = = 5e^{-2}$$

فاذن

$$P^*(x) = [1 - F(2)]^{-1} . P(x)$$

$$= [1 - 5e^{-2}]^{-1} . \frac{2^x e^{-2}}{x!}, x = 3, 4, 5, \ldots$$

لاحظ ان

$$\sum_{x=3}^{\infty} P^*(x) = e^{-2}(1 - 5e^{-2})^{-1} \sum_{x=3}^{\infty} \frac{2^x}{x!}$$

$$= e^{-2}(1 - 5e^{-2})^{-1} \cdot \left(\sum_{x=0}^{\infty} \frac{2^x}{x!} - 5 \right)$$

$$= e^{-2}(1 - 5e^{-2})^{-1}.(e^2 - 5) = (1 - 5e^{-2})^{-1}.(1 - 5e^{-2}) = 1$$

والشكل (٣ ـ ٨) يوضح مقارنة بين مخطط الدالة $P(x)$ والدالة $P^*(x)$

الشكل (٣ ـ ٨) ، مقارنة بين مخططي الدالتين $P^*(x), P(x)$

تمارين الفصل الثالث

٣ ـ ١ : ليكن X متغير عشوائي بدالة كثافة احتمالية

$$f(x) = cx(2-x)\,;\,0 < x \leq 2$$

جد : قيمة C ، المنوال ، الوسيط ، الربيع الثالث ، العشير الثامن ، الانحراف الربيعي ، معامل الاختلاف ، معامل التشتت ، معامل الالتواء

٣ ـ ٢ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = ce^{-b(x-a)}\,;\,x \geq a$
حيث ان a,b,c ثوابت حقيقية . يطلب اجراء مايلي :

أ ـ بين ان $c = b = \sigma_x^{-1}$ وان $a = \mu_x - \sigma_x$.

ب ـ جد المنوال والوسيط لقيم X في هذا التوزيع .

جـ ـ جد الربيعات والعشيرات لهذا التوزيع .

د ـ جد الانحراف الربيعي ، معامل التشتت .

هـ ـ ماهي قيمة ونوع الالتواء في هذا التوزيع؟

٣ ـ ٣ : لتكن $f(x) = b^{-2}xe^{-\frac{x^2}{2b^2}}, x > 0$ تمثل دالة الكثافة الاحتمالية للمتغير X . جد ما هو مطلوب في السؤال (٣ ـ ٢) عدا الفرع (أ) .

٣ ـ ٤ : ليكن X متغير عشوائي بدالة كثافة احتمالية $f(x) = c\sin x\,;\,0 \leq x \leq \frac{\pi}{2}$ جد قيمة الثابت c ثم جد ما هو مطلوب في السؤال (٣ ـ ٢) عدا الفرع (أ) .

٣ ـ ٥ : لتكن $f(x) = c\sin\frac{\pi x}{5}, 0 \leq x \leq 5$ تمثل دالة الكثافة الاحتمالية للمتغير X . جد قيمة الثابت C ثم استنتج دالة الكثافة الاحتمالية الى X معرفة قيمة على الفترات التالية : $[2,4], [1,5], [0,3]$.

٣ ـ ٦ : افرض ان X متغير عشوائي بدالة كثافة احتمالية f(x) وان

$P_r(X \leq x) = 1 - e^{-bx^2}, b > 0, x > 0$ جد مايلي :

أ ـ دالة الكثافة الاحتمالية للمتغير X .

ب ـ المنوال والوسيط في هذا التوزيع .

جـ ـ دالة الكثافة الاحتمالية الى X معرفة قيمه على الفترة $[3, \infty)$.

٣ ـ ٧ : افرض ان $P(x) = ab^x ; x = 0, 1, 2, \ldots$, $a + b = 1$. تمثل دالة الكتلة الاحتمالية للمتغير X . جد مايلي :

أ ـ المنوال والوسيط في هذا التوزيع .

ب ـ الربيع الاول والثالث في هذا التوزيع .

جـ ـ جد دالة الكتلة الاحتمالية الى X معرفة قيمه بالمجموعة A حيث

$A = \{ x : x = 4, 5, \ldots \}$

٣ ـ ٨ : برهن اذا كان التوزيع الاحتمالي للمتغير العشوائي المستمر X متماثلاً عندئذ

$$M = \frac{Q_1 + Q_3}{2} = \frac{D_i + D_{10-i}}{2} \quad , i = 1, 2, \ldots 9$$

حيث M تمثل الوسيط . Q_1, Q_3 : الربيع الاول والثالث . D_i : العشير i .

٣ ـ ٩ : اذا كان $S_k = \frac{\mu_X - M_0}{\sigma_x} = \frac{3(\mu_x - M_e)}{\sigma_x}$ حيث M_0 , M_e هما على التوالي الوسيط والمنوال في توزيع احتمالي لمتغير عشوائي مستمر مثل X . وبفرض ان التوزيع ذو التواء سالب . برهن ان $\mu_x < M_e < M_0$.

٣ ـ ١٠ : لتكن $f(x) = 6x(1 - x) ; 0 < x < 1$ تمثل دالة الكثافة الاحتمالية للمتغير X . جد دالة الكثافة الاحتمالية لهذا المتغير معرفة قيمه على الفترة $\left[\frac{1}{4}, 1 \right]$. للتوزيع الجديد جد : العزم ذا المرتبة r حول نقطة الاصل ، الوسط والتباين ، الدالة التوزيعية ، الربيع الثالث ، العشير السادس ، المنوال .





التوزيعات المشتركة ، الحدية ، الشرطية

الفصل الرابع
التوزيعات المشتركة ، الحدية والشرطية

يهدف هذا الفصل دراسة نوع آخر من الدوال الاحتمالية تتميز بكونها تتضمن اكثر من متغير واحد . على سبيل المثال دراسة العلاقة بين طول الفرد (متغير اول) . وزنه (متغير ثاني) وعمره (متغير ثالث) . العلاقة بين الانفاق على سلعة معينة (متغير اول) . دخل الفرد (متغير ثاني) . اسعار السلع البديلة (متغير ثالث) وعدد افراد العائلة (متغير رابع) . وغيرها من الامثلة الاخرى . وغالباً ما يسمى هذاالنوع من الدوال بدوال (توزيعات) متعددة المتغيرات Multivariate distributions . وسوف نركز الاهتمام في دراسة هذا النوع من الدوال من حيث خصائصها وعزومها وغيرها من الامور ذات العلاقة بها .

٤ ـ ١ : التوزيع المشترك Joint distribution

يعرف التوزيع المشترك بانه دالة احتمالية تجمع بين عدة متغيرات عشوائية في آن واحد . فعلى فرض ان X_1 , X_2 متغيران عشوائيان عندئذ فان النموذج الرياضي الاحتمالي الذي يعبر عن سلوك هذين المتغيرين معاً يسمى " التوزيع الاحتمالي المشترك " للمتغيرين X_1 , X_2 أو التوزيع الاحتمالي لمتغيرين Bivariate distribution . كما هي الحال لدى دراسة العلاقة بين الكمية المعروضة من سلعة معينة (X_1) وسعر الوحدة الواحدة منها (X_2) . وبشكل اكثر عمومية اذا كانت $X_1 , X_2 , \ldots , X_k$ متغيرات عشوائية عندئذ فان النموذج الرياضي الاحتمالي الذي يعبر عن سلوك **هذه المتغيرات** مجتمعة يسمى " التوزيع الاحتمالي المشترك " للمتغيرات $X_1 , X_2 , \ldots , X_k$ او التوزيع الاحتمالي لعدة متغيرات Multivariate distribution . ومن وجهة نظر احتمالية يمكن تحديد نوعين رئيسين من التوزيعات المشتركة استناداً الى نوع المتغيرات المتضمنة في التوزيع المشترك من حيث كونها متغيرات متقطعة ام مستمرة . هذين النوعين هما :

٤ ـ ١ ـ ١ : دوال الكتلة الاحتمالية المشتركة

Joint probability mass functions

افرض ان X_1 , X_2 متغيران عشوائيان من النوع المتقطع ، عندئذٍ فان دالة الكتلة الاحتمالية التي تعبر عن سلوك هذين المتغيرين معا هي

$$P(x_1 , x_2) = P_r(X_1 = x_1 , X_2 = x_2)$$

وهذا يعني ان

$$P(x_1 , x_2) = P_r(X_1 = x_1 \cap X_2 = x_2)$$

اي احتمال وقوع الحادثة

$$E = \{X_1 = x_1 \cap X_2 = x_2\}$$

وبشكل عام اذا كانت $X_1 , X_2 , \ldots , X_k$ متغيرات عشوائية من النوع المتقطع فاذن دالة الكتلة الاحتمالية التي تعبر عن سلوك هذه المتغيرات مجتمعة هي

$$P(x_1 , x_2 , \ldots , x_k) = P_r(X_1 = x_1 \cap X_2 = x_2 \cap \ldots \cap X_k = x_k)$$

أي احتمال وقوع الحادثة $E = \{X_1 = x_1 \cap X_2 = x_2 \cap \ldots \cap X_k = x_k\}$. ان دالة التوزيع المشترك يجب ان تحقق الشروط التالية التي تسمح لنا اعتبارها دالة كتلة احتمالية مشتركة وهي :

١ ـ ان دالة التوزيع المشترك دالة وحيدة القيمة عند اية قيمة مخصصة لمتغيرات التوزيع مثل $(x_1 , x_2 , \ldots , x_k)$

٢ ـ ان دالة التوزيع المشترك دالة غير سالبة كونها تعبر عن احتمال وقوع حادثة مثل E . أي ان $P(x_1 , x_2 , \ldots , x_k) \geq 0$

٣ ـ ان مجموع الكتل الاحتمالية المشتركة المقترنة بالقيم الممكنة الى $(X_1 , X_2 , \ldots , X_k)$ يجب ان يكون مساويا للواحد . أي ان :

$$\sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} P(x_1 , x_2 , \ldots , x_k) = 1$$

مثال (١) : لتكن $P(x_1 , x_2) = \frac{x_1 + x_2}{21}$, $x_1 = 1 , 2 , 3$, $x_2 = 1 , 2$ تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1 , X_2 . عندئذٍ .

$$P(X_1 = 1 , X_2 = 2) = \frac{1}{7} \quad , \quad P(X_1 = x_1 , X_2 = x_2) > 0$$

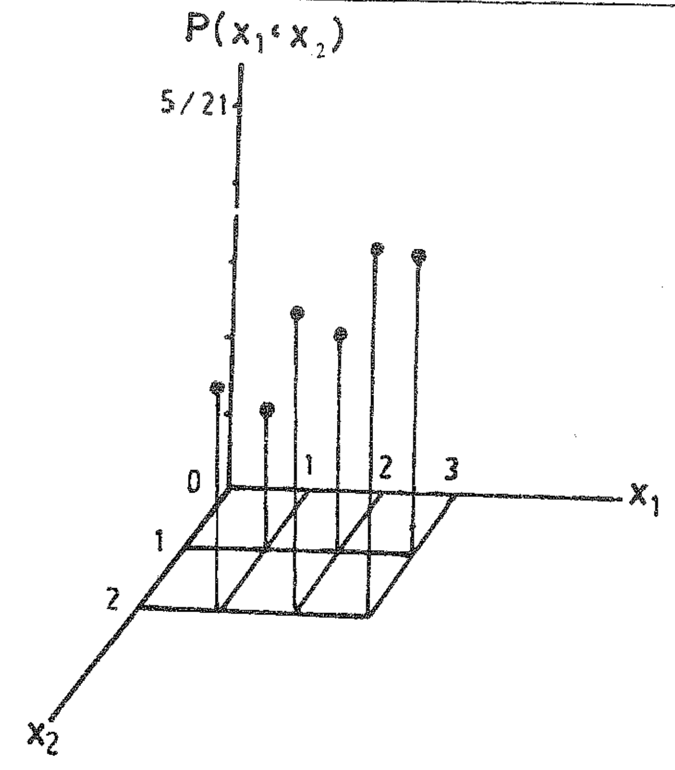
اي ان هذه الدالة غير سالبة وذات قيمة واحدة لاي زوج مثل (x_1, x_2) كذلك :

$$\sum_{x_1=1}^{3} \sum_{x_2=1}^{2} \left(\frac{x_1 + x_2}{21} \right) = \frac{1}{21} \sum_{x_1=1}^{3} (2x_1 + 3) = \frac{1}{21} . 21 = 1$$

لاحظ تحقق الشروط الثلاث التي سمحت لنا اعتبار $P(x_1, x_2)$ دالة كتلة احتمالية مشتركة بالمتغيرين X_1, X_2 . والاتي التوزيع الاحتمالي المشترك للمتغيرين (X_1, X_2) .

(x_1, x_2)	(1,1)	(1,2)	(2,1)	(2,2)	(3,1)	(3,2)
$P(x_1, x_2)$	2/21	3/21	3/12	4/21	4/21	5/21

والشكل (٤ ـ ١) يوضح مخطط التوزيع الاحتمالي للدالة $P(x_1, x_2)$



الشكل (٤ ـ ١) : مخطط الدالة $P(x_1, x_2)$.

٤ ـ ١ ـ ٢ : دوال الكثافة الاحتمالية المشتركة

Joint probability density functions

افرض ان X_1 , X_2 متغيران عشوائيان من النوع المستمر . عندئذ فان دالة الكثافة الاحتمالية المشتركة التي تعبر عن سلوك هذين المتغيرين معاً هي $f(x_1 , x_2)$.. ان $f(X_1=x_1,X_2=x_2)$ لا تعبرهنا عن قيمة احتمالية بل تعبر عن قيمة الدالة f عند قيمة مخصصة الى X_1 , X_2 مثل (x_1 , x_2) . وبشكل عام اذا كانت $X_1 , X_2 , \ldots , X_k$. متغيرات عشوائية من النوع المستمر فان الدالة $f(x_1 , x_2 , \ldots , x_k)$ تسمى دالة الكثافة الاحتمالية المشتركة التي تعبر عن سلوك المتغيرات $(X_1 , X_2 , \ldots , X_k)$ مجتمعة . وهذا النوع من الدوال يجب ان يحقق الخصائص التالية التي تسمح لنا اعتبار f دالة كثافة احتمالية مشتركة وهي :

١ ـ ان الدالة f دالة وحيدة القيمة عند أية قيمة مخصصة الى $(X_1 , X_2 , \ldots , X_k)$ مثل $(x_1 , x_2 , \ldots , x_k)$.

٢ ـ ان الدالة f دالة غير سالبة عند اية قيمة مخصصة الى $(X_1 , X_2 , \ldots , X_k)$ مثل $(x_1 , x_2 , \ldots , x_k)$.

٣ ـ انه وبشكل عام

$$\int_{x_1} \cdots \int_{x_{k-1}} \int_{x_k} f(x_1 , x_2 , \ldots , x_k)\, dx_k \,.\, dx_{k-1} \ldots dx_1 = 1$$

مثال (٢) : لتكن $f(x_1 , x_2)=e^{-(x_1+x_2)}, 0 < x_1 , x_2 < \infty$ تمثل الدالة المشتركة للمتغيرين X_1 , X_2 . لاحظ ان :

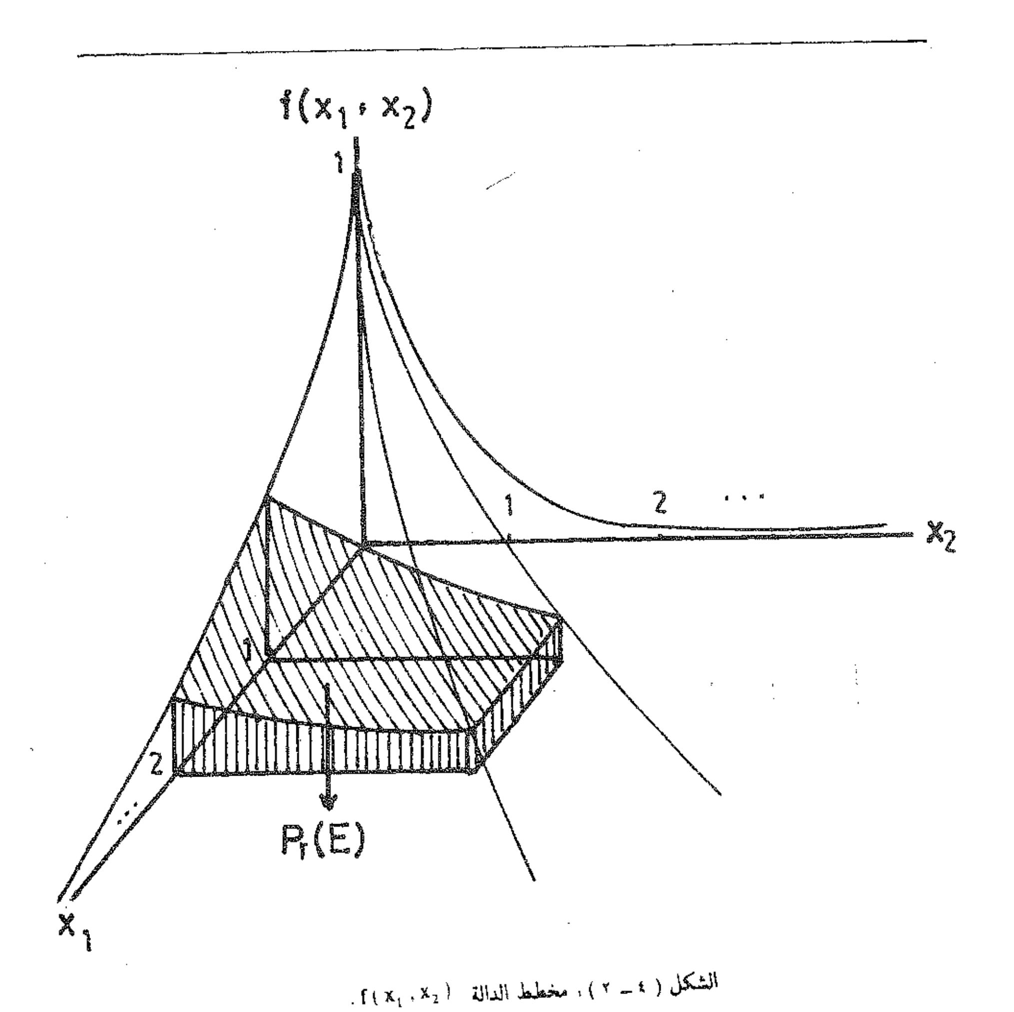
$$f(x_1 = 2 , x_2 = 3) = e^{-5} > 0$$

وان

$$\int_0^{\infty} \int_0^{\infty} e^{-(x_1+x_2)}\, dx_2 \,.\, dx_1 = \int_0^{\infty} \left[\int_0^{\infty} e^{-x_2} dx_2 \right] e^{-x_1}\, dx_1$$

$$= \int_0^{\infty} e^{-x_1}\, dx_1 = 1$$

وبتحقق الخصائص الثلاث الانفة الذكر فذلك يعني ان $f(x_1, x_2)$ هي دالة كثافة احتمالية مشتركة . والشكل (٤ ـ ٢) يوضح مخطط هذه الدالة .



الشكل (٤ ـ ٢) : مخطط الدالة $f(x_1, x_2)$.

واذا تطلب الامر حساب الاحتمال المشترك للحادثة $E=\{1< x_1 < 2 \cap 0< x_2 < 2\}$ أي حساب $P_r(1 < X_1 < 2, 0 < X_2 < 2)$ فان ذلك يتم وفق الآتي :

$$P_r(E) = \int_1^2 \int_0^2 e^{-(x_1+x_2)}\,dx_2 dx_1 = \int_1^2 \left[\int_0^2 e^{-x_2}\,dx_2 \right] e^{-x_1} dx_1$$

$$= (1 - e^{-2}) \int_1^2 e^{-x_1} dx_1 = (1 - e^{-2})(e^{-1} - e^{-2}) = 0{\cdot}2010727$$

وكما هو موضح بالجزء المظلل في الشكل (٤ ــ ٢) .

ملاحظة : قد يحدث في بعض الاحيان ان تكون بعض المتغيرات المتظمنة في الدالة المشتركة من النوع المتقطع والبقية من النوع المستمر . وسوف نشير لذلك لدى دراستنا لموضوع التوزيعات المركبة (او خلط التوزيعات) في الفقرة (٦ ــ ٧) .

٤ ــ ١ ــ ٣ : الدالة التوزيعية المشتركة Joint distribution function

تعرف الدالة التوزيعية المشتركة لتوزيع احتمالي مشترك بانها قيمة الاحتمال المتراكم لغاية قيمة معطاة الى $(X_1, X_2, \ldots, X_k)$ لتكن $(x_1, x_2, \ldots, x_k)$. فاذا رمزنا للدالة التوزيعية بالرمز $F(x_1, x_2, \ldots, x_k)$ عندئذٍ :

$$F(x_1, x_2, \ldots, x_k) = P_r(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_k \leq x_k)$$

فاذا كانت المتغيرات من النوع المتقطع فانه وبشكل عام :

$$F(x_1, x_2, \ldots, x_k) = \sum_{-\infty}^{x_1} \sum_{-\infty}^{x_2} \cdots \sum_{-\infty}^{x_k} P(u_1, u_2, \ldots, u_k)$$

واذا كان المتغيرات من النوع المستمر فانه وبشكل عام

$$F(x_1, x_2, \ldots, x_k) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_k} f(u_1, u_2, \ldots u_k) du_k . du_{k-1} \ldots du_1$$

وفي حالة المتغيرات المستمرة فان :

$$\frac{\partial^k F(x_1, x_2, \ldots, x_k)}{\partial x_1 . \partial x_2 \ldots \partial x_k} = f(x_1, x_2, \ldots, x_k)$$

ان خصائص الدالة التوزيعية . وبفرض وجود متغيرين في دالة مشتركة مثل X_1, X_2 . هي الاتي مع ملاحظة ان هذه الخصائص ستورد لحالة المتغيرات المستمرة وهي ذاتها صحيحة لحالة المتغيرات المتقطعة بمجرد استبدال رمز التكامل برمز الجمع اينما وجد :

١ ــ ان

$$F(-\infty, x_2) = \lim_{x_1 \to -\infty} F(x_1, x_2) = 0, F(x_1, -\infty) = \lim_{x_2 \to -\infty} F(x_1, x_2) = 0$$

وان

$$\lim_{\substack{x_1 \to \infty \\ x_2 \to \infty}} F(x_1, x_2) = 1 \quad , \lim_{\substack{x_1 \to -\infty \\ x_2 \to -\infty}} F(x_1, x_2) = 0$$

٢ ــ ان

$$\lim_{x_1 \to \infty} F(x_1, x_2) = F(x_2), \quad \lim_{x_2 \to \infty} F(x_1, x_2) = F(x_1)$$

البرهان :

$$F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(u_1, u_2)\, du_2 du_1$$

$$= \int_{-\infty}^{x_1} \left[\int_{-\infty}^{x_2} f(u_1, u_2)\, du_2 \right] du_1$$

فاذن

$$\lim_{x_2 \to \infty} F(x_1, x_2) = \int_{-\infty}^{x_1} \left[\int_{-\infty}^{\infty} f(u_1, u_2)\, du_2 \right] du_1$$

$$= \int_{-\infty}^{x_1} f(u_1)\, du_1 = F(x_1)$$

ووفق نفس الاسلوب يمكن برهان الحالة الاولى .

٣ ـ لتكن a_1, b_1, a_2, b_2 ثوابت حقيقية عندئذ :

$$P_r(a_1 < X_1 \le b_1, a_2 < X_2 \le b_2) = F(b_1, b_2) + F(a_1, a_2) - F(a_1, b_2) - F(b_1, a_2)$$

البرهان :

$$P_r(a_1 < X_1 \le b_1, a_2 < X_2 \le b_2) \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1, x_2) dx_2 dx_1 = P_r(E)$$

$$= \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} f(x_1, x_2) dx_2 dx_1 - \int_{-\infty}^{b_1} \int_{-\infty}^{a_2} f(x_1, x_2) dx_2 dx_1$$

$$- \int_{-\infty}^{b_2} \int_{-\infty}^{a_1} f(x_1, x_2) dx_2 dx_1 + \int_{-\infty}^{a_1} \int_{-\infty}^{a_2} f(x_1, x_2) dx_2 dx_1$$

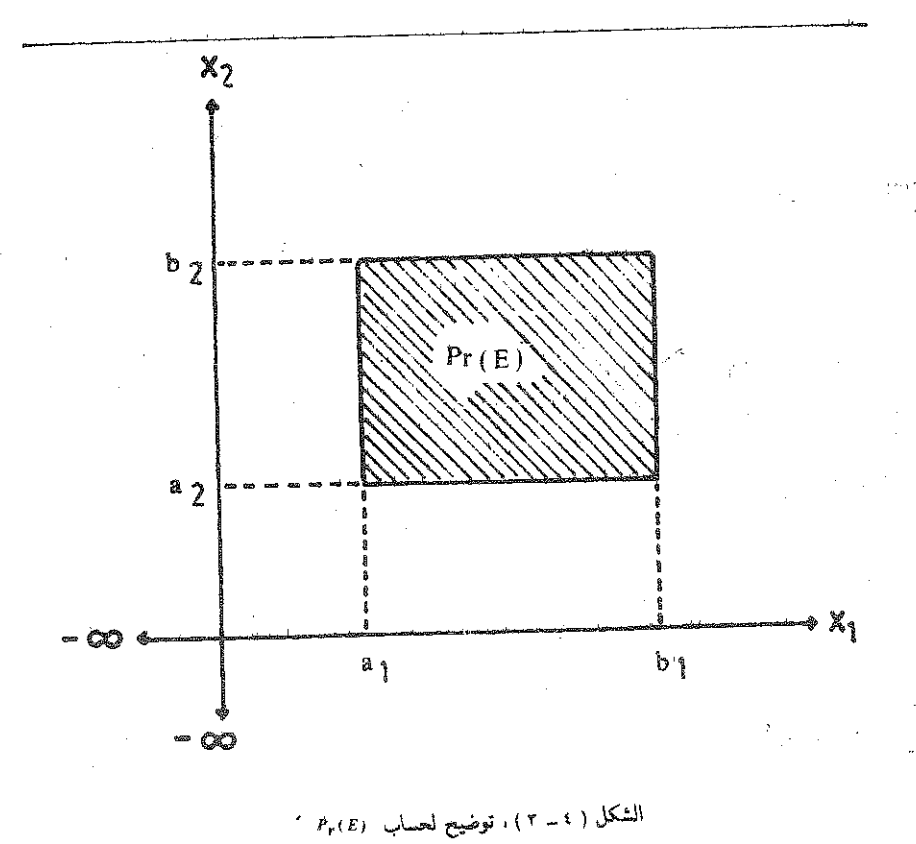
$$= F(b_1, b_2) - F(b_1, a_2) - F(b_2, a_1) + F(a_1, a_2)$$

وكما هو موضح في الشكل (٤ ـ ٣) .

٤ ـ ان $0 \le F(x_1, x_2) \le 1$ طالما ان F تمثل التراكم الاحتمالي لغاية قيمة معطاة او (X_1, X_2) مثل (x_1, x_2) .

٥ ـ ان الدالة $F(x_1, x_2)$ مستمرة نحو الجانب الايمن وغير متناقصة . فاذا كانت $b_1 \le b_2, a_1 \le a_2$ اعداد حقيقية فان .

$$F(a_1, b_1) \le F(a_2, b_1) \le F(a_2, b_2)$$



الشكل (٤ ــ ٢) ، توضيح لحساب $P_r(E)$

مثال (٣) : لتكن $P(x_1, x_2) = c \; ; x_1 = 1, 2, 3, 4, 5 \; ; x_2 = 1, 2, 3, 4$
تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2،جد قيمة c ثم جد الدالة التوزيعية المشتركة .

الحل : $\sum_{x_1=1}^{5} \sum_{x_2=1}^{4} P(x_1, x_2) = 1 \qquad \therefore \sum_{x_1=1}^{5} \sum_{x_2=1}^{4} c = 1$

فاذن

$$\sum_{x_1=1}^{5} 4c = 20c = 1 \qquad \therefore c = \frac{1}{20} \qquad \therefore P(x_1, x_2) = \frac{1}{20}$$

عليه فان

$$F(x_1, x_2) = \sum_{x_1=1}^{x_1} \sum_{x_2=1}^{x_2} \frac{1}{20} = \frac{x_1 x_2}{20}$$

$$\therefore F(x_1 x_2) = 0 \quad , x_1 < 1 \quad \text{or} \quad x_2 < 1$$

$$= \frac{x_1 x_2}{20} , 1 \le x_1 \le 5 , 1 \le x_2 \le 4$$

$$= 1 \quad , x_1 \ge 5 , x_2 \ge 4$$

لاحظ من هذه الدالة ما يلي :

ان $F(0, x_2) = F(x_1, 0) = 0$

ان $F(5, 4) = 1$

ان

$$P_r(3 < X_1 \le 5, 2 < X_2 \le 4) = F(5,4) + F(3,2) - F(3,4) - F(5,2)$$

$$= 1 + \frac{6}{20} - \frac{12}{20} - \frac{10}{20} = \frac{1}{5}$$

كذلك فان

$$F(3,5) = \frac{15}{20} > F(2,5) = \frac{10}{20} < F(2,4) = \frac{8}{20}$$

مثال (٤) : لتكن $f(x_1, x_2) = e^{-(x_1+x_2)} ; x_1, x_2 \ge 0$ تمثل دالة الكثافة الاحتمالية للمتغيرين X_1, X_2. جد الدالة التوزيعية المشتركة .

الحل :

$$F(x_1, x_2) = \int_0^{x_1} \int_0^{x_2} e^{-(u_1+u_2)} du_2 du_1$$

$$= \int_0^{x_1} \left[\int_0^{x_2} e^{-u_2} du_2 \right] e^{-u_1} du_1$$

$$= (1 - e^{-x_2}) \int_0^{x_1} e^{-u_1} du_1 = (1 - e^{-x_2})(1 - e^{-x_1})$$

فاذن

$$F(x_1, x_2) = 0, x_1 \leq 0 \quad \text{or} \quad x_2 \leq 0$$
$$= (1 - e^{-x_1})(1 - e^{-x_2}), 0 < x_1, x_2 < \infty$$
$$= 1, x_1 \to \infty, x_2 \to \infty$$

يلاحظ من هذه الدالة مايلي :

١ ـ ان

$$\frac{\partial^2 F(x_1, x_2)}{\partial x_1 . \partial x_2} = e^{-(x_1+x_2)} = f(x_1, x_2)$$

٢ ـ ان

$$F(0, x_2) = F(x_1, 0) = 0$$

٣ ـ ان

$$F(\infty, \infty) = \lim_{\substack{x_1 \to \infty \\ x_2 \to \infty}} (1 - e^{-x})(1 - e^{-x_2}) = 1$$

٤ ـ ان

$$P_r(1 < X_1 < 3, 2 < X_2 < 4) = F(3,4) + F(1,2) - F(1,4) - F(3,2)$$
$$= (1 - e^{-3})(1 - e^{-4}) + (1 - e^{-1})(1 - e^{-2}) - (1 - e^{-1})(1 - e^{-4})$$
$$- (1 - e^{-3})(1 - e^{-2})$$
$$= (e^{-1} - e^{-3})(e^{-2} - e^{-4}) = 0{\cdot}037223$$

٥ ـ ان

$$F(x_1, \infty) = F(x_1) = 1 - e^{-x_1}, F(\infty, x_2) = F(x_2) = 1 - e^{-x_2}$$

٤ ـ ١ ـ ٤ : التوقع الرياضي المشترك وتطبيقاته

Joint mathematical expectation and it's application

افرض ان $g(x_1, x_2, \ldots, x_k)$ دالة بدلالة المتغيرات العشوائية $X_1, X_2, \ldots, X_k$ عندئذ يعرف التوقع الرياضي للدالة g بانه « متوسط الدالة g » في ذلك التوزيع الاحتمالي المشترك . ويتم حساب التوقع الرياضي للدالة g في حالة المتغيرات المتقطعة وفق الآتي :

$$Eg(X_1, X_2, \ldots, X_k) = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} g(x_1, x_2, \ldots, x_k)\, P(x_1, x_2, \ldots, x_k)$$

اما في حالة المتغيرات المستمرة فان هذا التوقع يتم حسابه وبشكل عام وفق ما يلي :

$$Eg(X_1, X_2, \ldots, X_k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_k) \cdot f(x_1, x_2, \ldots, x_k)\, dx_k\, dx_{k-1} \ldots dx_1$$

وفي كلتا الحالتين يشترط . لتعريف التوقع الرياضي للدالة g . ان تكون عمليات الجمع او التكامل متقاربة على نحو مطلق اي ان :

$$\sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} |g(x_1, x_2, \ldots x_k)|\, P(x_1, x_2, \ldots, x_k) < \infty$$

وان

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |g(x_1, x_2, \ldots, x_k)|\, f(x_1, x_2, \ldots, x_k)\, dx_k dx_{k-1} \ldots dx_1 < \infty$$

وفي هذه الحالة يقال ان توقع الدالة g موجود .

وفيما يلي بعض خصائص التوقع الرياضي المشترك وتطبيقاته ،

١ ـ اذا كانت $g(x_1, x_2, \ldots, x_k) = c$ حيث c ثابت حقيقي فان $Eg = c$ ،

٢ ـ اذا كانت $g(x_1, x_2, \ldots, x_k) = x_i$ ، $i = 1, 2, \ldots, k$ فان

$$Eg(X_1, X_2, \ldots, X_k) = EX_i = \mu_i \, , i = 1, 2, \ldots, k$$

وهذا هو المتوسط لقيم المتغير X_i في ذلك التوزيع المشترك .

٣ ـ اذا كانت $g(x_1, x_2, \ldots, x_k) = (x_i - \mu_i)^2$ ، $i = 1, 2, \ldots, k$ فان :

$$Eg(X_1, X_2, \ldots, X_k) = E(X_i - \mu_i)^2 = V(X_i) = \sigma_i^2 \, , i = 1, 2, \ldots, k$$

وهذا هو التباين لقيم المتغير X_i في ذلك التوزيع المشترك .

٤ ـ اذا كانت $g(x_1, x_2, \ldots, x_k) = x_i \cdot x_j$ ، $i < j$ فان :

$$Eg(X_1, X_2, \ldots, X_k) = EX_i X_j = \mu_{ij} \, , i < j$$

وهذا هو العزم المشترك ذو المرتبة الثانية حول نقطة الأصل للمتغيرين X_i , X_j وهو في الحقيقة « متوسط حاصل ضرب المتغير X_i بالمتغير X_j »

٥ ـ اذا كانت $g(x_1, x_2, \ldots, x_k) = x_i \pm x_j$ فان :

$$Eg(X_1, X_2, \ldots, X_k) = E(X_i \pm X_j) = EX_i \pm EX_j = \mu_i \pm \mu_j$$

وبشكل عام اذا كانت $a_1, a_2, \ldots, a_k$ ثوابت حقيقية $X_1, X_2, \ldots, X_k$ متغيرات عشوائية وان $g(x_1, x_2, \ldots, x_k) = \sum_{i=1}^{k} a_i x_i$ وتمثل تركيب خطي Linear combination بدلالة هذه المتغيرات . عندئذٍ :

$$Eg(X_1, X_2, \ldots, X_k) = E\left[\sum_{i=1}^{k} a_i X_i\right]$$

$$= \sum_{i=1}^{k} E a_i X_i = \sum_{i=1}^{k} a_i EX_i$$

$$= \sum_{i=1}^{k} a_i \mu_i$$

مثال (٥) : افرض ان $p(x_1, x_2) = \frac{x_1 + x_2}{21}$; $x_1 = 1, 2, 3$; $x_2 = 1, 2$
تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 . عندئذ :

$$EX_1 = \sum_{x_1=1}^{3} \sum_{x_2=1}^{2} \left(\frac{x_1 + x_2}{21} \right)$$

$$= \frac{1}{21} \sum_{x_1=1}^{3} \left[\sum_{x_2=1}^{2} (x_1^2 + x_1 x_2) \right]$$

$$= \frac{1}{21} \sum_{x_1=1}^{3} (2x_1^2 + 3x_1) = \frac{46}{21}$$

كذلك فان

$$EX_2 = \sum_{x_1=1}^{3} \sum_{x_2=1}^{2} x_2 \left(\frac{x_1 + x_2}{21} \right) = \frac{11}{7}$$

وان

$$EX_1 X_2 = \sum_{x_1=1}^{3} \sum_{x_2=1}^{2} x_1 x_1 \left(\frac{x_1 + x_2}{21} \right)$$

$$= \frac{1}{21} \sum_{x_1=1}^{3} \left[\sum_{x_2=1}^{2} x_1 x_2 (x_1 + x_2) \right]$$

$$= \frac{1}{21} \sum_{x_1=1}^{3} \left[\sum_{x_2=1}^{2} (x_1^2 x_2 + x_1 x_2^2) \right]$$

$$= \frac{1}{21} \sum_{x_1=1}^{3} (3x_1^2 + 5x_1)$$

$$= \frac{1}{21} (42 + 30) = \frac{72}{21}$$

كذلك فان

$$E(2X_1 + 3X_2) = 2EX_1 + 3EX_2 = 2 \left(\frac{46}{21} \right) + 3 \left(\frac{11}{7} \right) = \frac{191}{21}$$

وان

$$E(4X_1 - 2X_2) = 4EX_1 - 2EX_2 = 4 \left(\frac{46}{21} \right) - 2 \left(\frac{11}{7} \right)$$

$$= \frac{118}{21}$$

مثال (٦) : لتكن $f(x_1, x_2) = x_1 + x_2 ; 0 < x_1, x_2 < 1$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1, X_2 . عندئذ :

$$EX_1 = \int_0^1 \int_0^1 x_1 (x_1 + x_2)\, dx_2\, dx_1$$

$$= \int_0^1 \int_0^1 x_1^2\, dx_2\, dx_1 + \int_0^1 \int_0^1 x_1 x_2\, dx_2\, dx_1$$

$$= \int_0^1 x_1^2\, dx_1 \cdot \int_0^1 dx_2 + \int_0^1 x_1\, dx_1 \cdot \int_0^1 x_2\, dx_2$$

$$= \left[\frac{x_1^3}{3} \right]_0^1 \cdot \left[x_2 \right]_0^1 + \left[\frac{x_1^2}{2} \right]_0^1 \cdot \left[\frac{x_2^2}{2} \right]_0^1$$

$$= \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

ووفق نفس الاسلوب يمكن البيان ان $EX_2 = \frac{7}{12}$ كذلك فان :

$$E X_1 X_2 = \int_0^1 \int_0^1 x_1 x_2 (x_1 + x_2)\, dx_2\, dx_1$$

$$= \int_0^1 \int_0^1 x_1^2 x_2\, dx_2\, dx_1 + \int_0^1 \int_0^1 x_1 x_2^2\, dx_2\, dx_1$$

$$= \int_0^1 x_1^2\, dx_1 \cdot \int_0^1 x_2\, dx_2 + \int_0^1 x_1\, dx_1 \cdot \int_0^1 x_2^2\, dx_2$$

$$= \left[\frac{x_1^3}{3} \right]_0^1 \cdot \left[\frac{x_2^2}{2} \right]_0^1 + \left[\frac{x_1^2}{2} \right]_0^1 \cdot \left[\frac{x_2^3}{3} \right]_0^1$$

$$= \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{3}$$

وان

$$E(2X_1 + 4X_2) = 2EX_1 + 4EX_2 = 2\left(\frac{7}{12}\right) + 4\left(\frac{7}{12}\right)$$
$$= 3.5$$

وان

$$E(3X_1 - 5X_2) = 3EX_1 - 5EX_2 = 3\left(\frac{7}{12}\right) - 5\left(\frac{7}{12}\right)$$
$$= -\frac{7}{6}$$

٤ ــ ١ ــ ٥ : التباين المشترك ومعاملات الارتباط
Covariance and correlation coefficients

اذا تم اختيار الدالة $g(x_1, x_2, \ldots, x_k) = (x_i - EX_i)(x_j - EX_j)$ فان :

$$Eg(X_1, X_2, \ldots, X_k) = E(X_i - EX_i)(X_j - EX_j) = \sigma_{ij}$$

وهذا غالباً ما يسمى « التباين المشترك » بين المتغيرين X_i, X_j . ويلاحظ ان

$$\text{cov}(X_i, X_j) = \sigma_{ij} = EX_i X_j - EX_i . EX_j$$

كذلك يمكن البيان وبسهولة ان $V(X_i \pm X_j) = V(X_i) + V(X_j) \pm \text{Cov}(X_i, X_j)$ ونترك برهنة ذلك للقاريء. وباستخدام جبر المصفوفات **matrix algebra** وبفرض ان X متجه vector ذا بُعد k عناصره تمثل مجموعة المتغيرات $X_1, X_2, \ldots, X_k$ اي ان $X = [X_1 \; X_2 \ldots X_k]$ يسلك وفق دالة كتلة او كثافة احتمالية مشتركة فان $EX = [EX_1 \; EX_2 \ldots EX_k]$ يمثل متجه المتوسطات الى X وان التباينات لهذه المتغيرات وكذلك التباينات المشتركة ما بين اي متغيرين منها معرفة في المصفوفة التالية التي تسمى « مصفوفة التباين والتباين المشترك للمتجه X » :

$$\text{Var} - \text{cov}(X) = \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ \sigma_{k1} & \sigma_{k2} & \sigma_{k3} & \cdots & \sigma_k^2 \end{bmatrix}$$

ان المصفوفة Σ مصفوفة متماثلة symmetric matrix ذات مرتبة KXK عناصر القطر الرئيسي فيها تمثل تباينات عناصر X . في حين ان العناصر الواقعة خارج القطر الرئيسي تمثل التباينات المشتركة بين اي عنصرين من عناصر X .

ومن خلال هذه المصفوفة يمكن الحصول على معامل الارتباط البسيط بين اي متغيرين من متغيرات المتجه X . فاذا رمزنا لمعامل الارتباط البسيط بين المتغيرين X_i , X_j بالرمز ρ_{ij} عندئذٍ يمكن حساب قيمة هذا المعامل وفق الصيغة التالية :

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \cdot \sigma_j} \; , \; -1 \leq \rho_{ij} \leq 1$$

وهذا يعني انه يمكن تعريف مصفوفة معاملات الارتباط البسيطة بين متغيرات المتجه X والمعطاة بالمصفوفة التالية التي تسمى « مصفوفة الارتباطات البسيطة » :

$$\mathbf{R} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & \rho_{k3} & \cdots & 1 \end{bmatrix}$$

وهي مصفوفة متماثلة ذات مرتبة KXK عناصر القطر الرئيسي فيها مساوية للواحد دلالة على ارتباط المتغير مع ذاته في حين ان العناصر خارج القطر الرئيسي تمثل معاملات الارتباط البسيطة ما بين عناصر المتجه X . ان معامل الارتباط البسيط مقياس لدرجة العلاقة بين متغيرين خالٍ من وحدات القياس . وكلما كانت $|\rho_{ij}|$ قريبة من الواحد فذلك مؤشر على قوة العلاقة بين X_i , X_j في حين انه كلما كانت $|\rho_{ij}|$ قريبة من الصفر فذلك مؤشر على ضعف العلاقة بينهما . اما اشارة هذا المعامل فانها تعني اتجاه العلاقة بين X_i , X_j فاذا كانت الاشارة موجبة (بسبب ان $\sigma_{ij} > 0$) فذلك يعني ان العلاقة موجبة (طردية) واذا كانت الاشارة سالبة (بسبب ان $\sigma_{ij} < 0$) فذلك يعني ان العلاقة سالبة (عكسية) .

قد يتطلب الامر في بعض الاحيان حساب درجة العلاقة بين متغيرين مثل X_i , X_j بعد استبعاد اثر متغير ثالث مثل X_ℓ مرتبط مع كل من X_j و X_i . ان معامل

الارتباط الذي يقيس علاقة من هذا النوع يسمى « معامل الارتباط الجزئي partial correlation coefficient » معطاة صيغته بما يلي :

$$\rho_{ij\cdot L} = \frac{\rho_{ij} - \rho_{iL}\cdot\rho_{jL}}{\sqrt{(1-\rho^2_{iL})(1-\rho^2_{jL})}} \quad ; |\rho_{iL}| \neq 1 , \quad |\rho_{jL}| \neq 1$$

وفي حالة وجود اربعة متغيرات مثل X_i , X_j , X_L , X_m فان صيغة معامل الارتباط الجزئي مابين X_i , X_j باستبعاد اثر X_L , X_m هي :

$$\rho_{ij\cdot Lm} = \frac{\rho_{ij\cdot L} - \rho_{im\cdot L}\cdot\rho_{im\cdot L}}{\sqrt{(1-\rho^2_{im\cdot L})\cdot(1-\rho^2_{jm\cdot L})}} \quad ; |\rho_{im\cdot L}| \neq 1 , |\rho_{jm\cdot L}| \neq 1$$

كذلك يتطلب الامر في بعض الاحيان حساب درجة العلاقة بين متغير واحد من جهة وعدة متغيرات من جهة أخرى . إن معامل الارتباط الذي يقيس علاقة من هذا النوع يسمى «معامل الإرتباط المتعدد multiple correlation coefficient معطاة صيغته في حالة وجود ثلاث متغيرات مثل X_i , X_j , X_L , وتطلب الامر حساب معامل الارتباط المتعدد بين X_i من جهة والمتغيرين X_j , X_L من جهة اخرى . بالآتي :

$$R_{i\cdot jL} = \sqrt{1-(1-\rho^2_{ij})(1-\rho^2_{iL\cdot j})}$$

وصيغة هذا المعامل في حالة وجود اربعة متغيرات هي :

$$R_{i.jLm} = \sqrt{1-(1-\rho^2_{ij})(1-\rho^2_{iL\cdot j})(1-\rho^2_{im\cdot jL})}$$

مثال (٧) : لمعطيات المثال (٥) . جد معامل الارتباط البسيط بين X_2 , X_1

الحل : $EX_1 = \frac{46}{21} , EX_2 = \frac{11}{7} , EX_1 X_2 = \frac{72}{21}$

فاذن $\sigma_{12} = EX_1 X_2 - EX_1 . EX_2 = -\frac{2}{147}$

وان

$$EX_1^2 = \sum_{x_1=1}^{3} \sum_{x_2=1}^{2} x_1^2 \left(\frac{x_1 + x_2}{21} \right) = \frac{114}{21}$$

$$\therefore \quad \sigma_1^2 = EX_1^2 - (EX_1)^2 = \frac{278}{441}$$

كذلك فان

$$EX_2^2 = \sum_{x_1=1}^{3} \sum_{x_2=1}^{2} x_2^2 \left(\frac{x_1 + x_2}{21} \right) = \frac{19}{7}$$

$$\therefore \quad \sigma_2^2 = EX_2^2 - (EX_2)^2 = \frac{12}{49}$$

فاذن :

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \cdot \sigma_2} = -0.035$$

وهذا يعني ان العلاقة بين X_1 ، X_2 ضعيفة وعكسية .

مثال (٨) : لمعطيات المثال (٦) جد معامل الارتباط البسيط بين X_1 و X_2 .

الحل :

$$EX_1 = EX_2 = \frac{7}{12}, \; EX_1X_2 = \frac{1}{3}$$

فاذن

$$\sigma_{12} = EX_1X_2 - EX_1 . EX_2 = -\frac{1}{144}$$

وان

$$EX_1^2 = \int_0^1 \int_0^1 x_1^2 (x_1 + x_2)\, dx_2 dx_1 = \frac{5}{12} = EX_2^2$$

فاذن

$$\sigma_1^2 = EX_1^2 - (EX_1)^2 = \frac{11}{144} = \sigma_2^2$$

عليه فان

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \cdot \sigma_2} = -\frac{1}{11}$$

٤ ـ ١ ـ ٦ : الدالة المولدة لعزوم التوزيعات المشتركة Joint moment generating function

افرض ان $X_1, X_2, ..., X_k$ متغيرات عشوائية وان $t_1, t_2, ..., t_k$ متغيرات اخرى وان h_i عدد موجب بحيث ان $-h_i < t_i < h_i$. عندئذٍ تعرف الدالة المولدة لعزوم التوزيع المشترك للمتغيرات $X_1, X_2, ..., X_k$ على النحو الآتي :

$$M(t_1, t_2, ..., t_k) = E\, e^{\sum_{i=1}^{k} t_i X_i}$$

بشرط ان التوقيع موجود .

فاذا كانت المتغيرات $X_1, X_2, ..., X_k$ من النوع المتقطع فان

$$M(t_1, t_2, ..., t_k) = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} e^{\sum_{i=1}^{k} t_i x_i} \cdot P(x_1, x_2, ..., x_k)$$

اما اذا كانت هذه المتغيرات من النوع المستمر فانه وبشكل عام :

$$M(t_1, t_2, ..., t_k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\sum_{i=1}^{k} t_i x_i} . f(x_1, x_2, ..., x_k) dx_k . dx_{k-1} \ldots dx_1$$

واضح من تعريف هذه الدالة ان :

$$M(t_1 = 0, t_2 = 0, \ldots, t_i = 0, \ldots, t_k = 0) = Ee^0 = 1$$

وان

$$M(t_1 = 0, t_2 = 0, \ldots, t_i \neq 0, t_{i+1} = 0, \ldots, t_k = 0) = Ee^{t_i X_i} = M_{X_i}(t_i)$$

وكما هو معلوم فان الهدف من هذا النوع من الدوال هو توليد عزوم التوزيع . وسوف نوضح هذه العملية في حالة وجود توزيع مشترك بمتغيرين . واضح هنا ان :

$$M(t_1, t_2) = Ee^{t_1 X_1 + t_2 X_2}$$

وباستخدام مفكوك سلسلة مكلورين فان :

$$e^{t_1 X_1 + t_2 X_2} = \sum_{u=0}^{\infty} \frac{(t_1 X_1 + t_2 X_2)^u}{u_i}$$

وهذا يعني ان :

$$M(t_1, t_2) = E \sum_{u=0}^{\infty} \frac{(t_1 X_1 + t_2 X_2)^u}{u!}$$

فاذن :

$$M(t_1, t_2) = E\left[1 + (t_1 X_1 + t_2 X_2) + \frac{(t_1 X_1 + t_2 X_2)^2}{2!} + \frac{(t_1 X_1 + t_2 X_2)^3}{3!} + \ldots\right]$$

$$= 1 + t_1 EX_1 + t_2 EX_2 + \frac{t_1^2}{2!} EX_1^2 + \frac{t_2^2}{2!} EX_2^2$$

$$+ t_1 t_2 EX_1 X_2 + \frac{t_1^3}{3!} EX_1^3 + \frac{3t_1^2 t_2}{3!} EX_1^2 X_2 + \frac{3t_1 t_2^2}{3!} EX_1 X_2^2$$

$$= \frac{t_2^3}{3!} EX_2^3 + \ldots$$

ويلاحظ من هذه الصيغة ان عزوم كل متغير حول نقطة الاصل موجودة وكذلك العزوم المشتركة . وهذا يعني انه يمكن « توليد » هذه العزوم لكل متغير بشكل منفرد وكذلك العزوم المشتركة ما بين X_1, X_2 وعلى النحو التالي :

بايجاد المشتقة الجزئية الاولى للدالة $M(t_1, t_2)$ نسبة الى t_1 نحصل على :

$$\frac{\partial M(t_1, t_2)}{\partial t_1} = EX_1 + t_1 EX_1^2 + t_2 EX_1X_2 + O'(t_1, t_2)$$

حيث $O'(t_1, t_2)$ تعني حدود لاحقة تمثل مشتقات جزئية من المرتبة الاولى تتضمن t_1 او t_2 او كليهما بقوى عليا . وبجعل $t_1 = t_2 = 0$ نحصل على :

$$\left.\frac{\partial M(t_1, t_2)}{\partial t_1}\right]_{t_1 = t_2 = 0} = EX_1$$

وهذا ماهو الا متوسط X_1 في التوزيع المشترك .

كذلك فان

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_1^2} = EX_1^2 + t_1 EX_1^3 + t_2 EX_1^2 X_2 + O''(t_1, t_2)$$

حيث $O''(t_1, t_2)$ نعني حدود لاحقة تمثل مشتقات جزئية من المرتبة الثانية تتضمن t_1 او t_2 او كليهما بقوى عليا . وبجعل $t_1 = t_2 = 0$ نحصل على :

$$\left.\frac{\partial^2 M(t_1, t_2)}{\partial t_1^2}\right]_{t_1 = t_2 = 0} = EX_1^2$$

وهذا ماهو الا العزم الثاني للمتغير X_1 حول نقطة الاصل . ووفق ماتقدم يمكن ملاحظة ان العزم ذو المرتبة r حول نقطة الاصل لاي من هذين المتغيرين ماهو الا :

$$EX_1^r = \left.\frac{\partial^r M(t_1, t_2)}{\partial t_1^r}\right]_{t_1 = t_2 = 0}, \quad EX_2^r = \left.\frac{\partial^r M(t_1, t_2)}{\partial t_2^2}\right]_{t_1 = t_2 = 0}$$

الان لوعدنا للمشتقة الجزئية الاولى نسبة الى t_1 وقمنا باشتقاقها نسبة الى t_2 فاننا نحصل على :

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_1 \partial t_2} = EX_1X_2 + t_1EX_1^2X_2 + t_2EX_1X_2^2 + O''(t_1, t_2)$$

وبجعل $t_1 = t_2 = 0$ نحصل على :

$$\left.\frac{\partial^2 M(t_1, t_2)}{\partial t_1 \partial t_2}\right]_{t_1 = t_2 = 0} = EX_1X_2$$

وهذا ماهو الا العزم المشترك ذو المرتبة الثانية حول نقطة الاصل . ووفق نفس المفهوم يمكن البيان ان :

$$\left.\frac{\partial^3 M(t_1, t_2)}{\partial t_1^2 \partial t_2}\right]_{t_1 = t_2 = 0} = EX_1^2X_2 , \quad \left.\frac{\partial^3 M(t_1, t_2)}{\partial t_1 \partial t_2^2}\right]_{t_1 = t_2 = 0} = EX_1X_2^2$$

عليه وبشكل عام اذا كان r_2 , r_1 عددين صحيحين فان :

$$\left.\frac{\partial^{r_1 + r_2} M(t_1, t_2)}{\partial^{r_1} t_1 \partial^{r_2} t_2}\right]_{t_1 = t_2 = 0} = EX_1^{r_1}. X_2^{r_2}.$$

وهذا يسمى العزم المشترك ذو المرتبة $(r_1 + r_2)$ حول نقطة الاصل .

ومما تقدم يمكن ملاحظة مايلي :

ان : $$\left[\frac{\partial^2 M(t_1, t_2)}{\partial t_i^2} - \left(\frac{\partial M(t_1, t_2)}{\partial t_i} \right)^2 \right]_{t_1 = t_2 = 0} = \sigma_i^2 , i = 1, 2$$

وان

$$\left[\frac{\partial^2 M(t_1, t_2)}{\partial t_1 \partial t_2} - \left(\frac{\partial M(t_1, t_2)}{\partial t_1} \right) \left(\frac{\partial M(t_1, t_2)}{\partial t_2} \right) \right]_{t_1 = t_2 = 0}$$

$$= \sigma_{12}$$

كذلك يمكن ايجاد دوال من شأنها توليد عزوم مشتركة مركزية . ففي حالة وجود متغيرين مثل X_1, X_2 فان الدالة المولدة للعزوم المشتركة المركزية تعرف بالشكل التالي :

$$M_c(t_1, t_2) = E\, e^{t_1(X_1 - EX_1) + t_2(X_2 - EX_2)}$$

حيث M_c تعني الدالة المولدة للعزوم المشتركة المركزية . وهذا يعني ان

$$M_c(t_1, t_2) = e^{-(t_1 EX_1 + t_2 EX_2)} . M(t_1, t_2)$$

ووفق نفس الاسس التي تم اعتمادها بشأن توليد العزوم المشتركة حول نقطة الاصل يمكن اعتمادها ايضاً في توليد العزوم المشتركة المركزية . حيث يمكن البيان ان :

$$\left.\frac{\partial^r M_c(t_1, t_2)}{\partial t_i^r}\right]_{t_1 = t_2 = 0} = E(X_i - EX_i)^r, i = 1, 2$$

التي منها يتبين ان

$$E(X_i - EX_i) = 0, E(X_i - EX_i)^2 = \sigma_i^2$$

كذلك فان :

$$\left.\frac{\partial^{r_1 + r_2} M_c(t_1, t_2)}{\partial t_1^{r_1} . \partial t_2^{r_2}}\right]_{t_1 = t_2 = 0} = E(X_1 - EX_1)^{r_1} . (X_2 - EX_2)^{r_2}$$

ويتضح من هذه الصيغة ان :

$$E(X_1 - EX_1)(X_2 - EX_2) = \sigma_{12}$$

مثال (٩) : افرض ان

$$P(x_1, x_2) = \frac{6!}{x_1! x_2!}\left(\frac{1}{4}\right)^{x_1} \cdot \left(\frac{3}{4}\right)^{x_2}, x_1 = 0, 1, \ldots, 6$$

$$x_2 = 6 - x_1$$

تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1 , X_2 . عندئذ فان :

$$M(t_1, t_2) = \sum_x e^{t_1 x_1 + t_2 x_2} \cdot \frac{6!}{x_1! \, x_2!} \left(\frac{1}{4}\right)^{x_1} \cdot \left(\frac{3}{4}\right)^{x_2}$$

$$= \sum_x \frac{6!}{x_1! \, x_2!} \left(\frac{1}{4} e^{t_1}\right)^{x_1} \cdot \left(\frac{3}{4} e^{t_2}\right)^{x_2}$$

$$= \sum_{x_1=0}^{6} \frac{6!}{x_1! \, (6 - x_1)!} \left(\frac{1}{4} e^{t_1}\right)^{x_1} \cdot \left(\frac{3}{4} e^{t_2}\right)^{6 - x_1}$$

$$= \sum_{x_1=0}^{6} C^6_{x_1} \left(\frac{1}{4} e^{t_1}\right)^{x_1} \cdot \left(\frac{3}{4} e^{t_2}\right)^{6 - x_1}$$

الان لاي عددين مثل a , b فانه باستخدام نظرية ثنائي الحدين يمكن البيان ان $(a + b)^6 = \sum_{k=0}^{6} C^6_k a^k b^{6-k}$. وهذا يعني انه لو تم جعل $a = \frac{1}{4} e^{t_1}$, $b = \frac{3}{4} e^{t_2}$, $k = x_1$ فان

$$\left(\frac{1}{4} e^{t_1} + \frac{3}{4} e^{t_2}\right)^6 = \sum_{x_1=0}^{6} C^6_{x_1} \left(\frac{1}{4} e^{t_1}\right)^{x_1} \left(\frac{3}{4} e^{t_2}\right)^{x_2}$$

من ذلك نستنتج ان :

$$M(t_1, t_2) = \left(\frac{1}{4} e^{t_1} + \frac{3}{4} e^{t_2}\right)^6$$

لاحظ ان

$$M(0, 0) = \left(\frac{1}{4} + \frac{3}{4}\right)^6 = 1$$

وان

$$M(0, t_2) = \left(\frac{1}{4} + \frac{3}{4} e^{t_2}\right)^6 , \quad M(t_1, 0) = \left(\frac{1}{4} e^{t_1} + \frac{3}{4}\right)^6$$

كذلك فان .

$$\frac{\partial M(t_1, t_2)}{\partial t_1} = \frac{6}{4} e^{t_1} \left(\frac{1}{4} e^{t_1} + \frac{3}{4} e^{t_2}\right)^5$$

وهذا يعني ان $\frac{3}{2} = EX_1$ وان

$$\frac{\partial M(t_1, t_2)}{\partial t_2} = \frac{18}{4} e^{t_2} \left(\frac{1}{4} e^{t_1} + \frac{3}{4} e^{t_2} \right)^5$$

وهذا يعني ان $\frac{9}{2} = EX_2$ كذلك فان

$$\left.\frac{\partial^2 M(t_1, t_2)}{\partial t_1^2}\right]_{t_1 = t_2 = 0} = \frac{27}{8}, \quad \left.\frac{\partial^2 M(t_1, t_2)}{\partial t_2^2}\right]_{t_1 = t_2 = 0} = \frac{171}{8}$$

وان

$$\sigma_1^2 = \frac{27}{8} - \left(\frac{3}{2} \right)^2 = \frac{9}{8}, \sigma_2^2 = \frac{171}{8} - \left(\frac{9}{2} \right)^2 = \frac{9}{8}$$

كذلك فان

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_1 \partial t_2} = \frac{45}{8} e^{t_1} . e^{t_2} \left(\frac{1}{4} e^{t_1} + \frac{3}{4} e^{t_2} \right)^4$$

وهذا يعني ان

$$\left.\frac{\partial^2 M(t_1, t_2)}{\partial t_1 . \partial t_2}\right]_{t_1 = t_2 = 0} = EX_1X_2 = \frac{45}{8}$$

فاذن

$$\sigma_{12} = \frac{45}{8} - \left(\frac{3}{2} \right) \left(\frac{9}{2} \right) = - \frac{9}{8}$$

عليه فان

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 . \sigma_2} = -1$$

كذلك فان

$$M_c(t_1, t_2) = e^{-\left(\frac{3}{2} t_1 + \frac{9}{2} t_2 \right)} . \left(\frac{1}{4} e^{t_1} + \frac{3}{4} e^{t_2} \right)^6$$

مثال (١٠) : افرض ان $f(x_1, x_2) = e^{-(x_1 + x_2)}; x_1, x_2 \geq 0$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1, X_2 . عندئذٍ فان :

$$M(t_1, t_2) = \int_0^\infty \int_0^\infty e^{t_1 x_1 + t_2 x_2} \cdot e^{-(x_1 + x_2)} \, dx_2 dx_1$$

$$= \int_0^\infty \int_0^\infty e^{-x_1(1-t_1)} \cdot e^{-x_2(1-t_2} \, dx_2 dx_1$$

$$= \int_0^\infty e^{-x_1(1-t_1)} \, dx_1 \cdot \int_0^\infty e^{-x_2(1-t_2)} \, dx_2$$

$$= \frac{1}{(1-t_1)(1-t_2)} \; ; t_1, t_2 < 1$$

لاحظ من هذا المثال ما يلي :

ان $M(0,0) = 1$. وان :

$$\left.\frac{\partial M(t_1, t_2)}{\partial t_1}\right]_{t_1 = t_2 = 0} = [(1-t_1)^2 \cdot (1-t_2)^{-1}]_{t_1 = t_2 = 0} = 1 = 1!$$

$$\left.\frac{\partial^2 M(t_1, t_2)}{\partial t_1^2}\right]_{t_1 = t_2 = 0} = 2[(1-t_1)^3 (1-t_2)^{-1}]_{t_1 = t_2 = 0} = 2 = 2!$$

$$\left.\frac{\partial^2 M(t_1, t_2)}{\partial t_1^3}\right]_{t_1 = t_2 = 0} = 6[(1-t_1)^4 (1-t_2)^{-1}]_{t_1 = t_2 = 0} = 6 = 3!$$

وهذا يعني ان

$$\left.\frac{\partial^r M(t_1, t_2)}{\partial t_1^r}\right]_{t_1 = t_2 = 0} = \left.\frac{\partial^r M(t_1, t_2)}{\partial t_2^r}\right]_{t_1 = t_2 = 0} = r!$$

ويترك للقاريء البيان ان

$$\left.\frac{\partial^{r_1 + r_2} M(t_1, t_2)}{\partial t_1^{r_1} \cdot \partial t_2^{r_2}}\right]_{t_1 = t_2 = 0} = (r_1 !)(r_2 !)$$

كذلك فان

$$M_c(t_1, t_2) = e^{-(t_1+t_2)} . [(1-t_1)(1-t_2)]^{-1} ; t_1, t_2 < 1$$

ويطلب من القاري حساب معامل الارتباط البسيط بين X_1, X_2 .

وفيما يلي بعض الملاحظات عن الدوال المولدة لعزوم التوزيعات المشتركة :

١ ـ ان خصائص الدوال المولدة لعزوم التوزيعات المشتركة من حيث المفهوم هي نفس الخصائص المنوه عنها في الفقرة (٢ ـ ٢ ـ ١)

٢ ـ قد تكون الدالة المولدة لعزوم توزيع مشترك موجودة وقد تكون غير موجودة وذلك يتوقف على امكانية تحديد توقع الدالة $e^{\Sigma t_i X_i}$

٣ ـ ان الدالة المولدة لعزوم توزيع مشترك (اذا كانت موجودة) هي دالة وحيدة تشخص التوزيع الاحتمالي المشترك الذي اشتقت منه .

٤ ـ ان كل توزيع مشترك يمتلك دالة مميزة معرفة بالشكل

$$\phi(t_1, t_2, \ldots, t_k) = Ee^{i\Sigma t_j X_j} \quad , i = \sqrt{-1}, j = 1, 2, \ldots, k$$

وهذا يعني ان هذه الدالة موجودة دائماً (بعكس الدالة المولدة للعزوم) بسبب ان $1 \leq |\phi(t_1, t_2, \ldots, t_k)|$. ويمكن ملاحظة ذلك حسب ماهو موضح في الفقرة (٢ ـ ٣ ـ ١) . وعن طريق الدالة المميزة يمكن ايضا استنتاج عزوم التوزيع المشترك .

٤ ـ ٢ : التوزيع الحدي Marginal distribution

استعرضنا في الفقرة (٤ ـ ١) مفهوم التوزيع المشترك وخصائصه واهم العزوم والمقاييس المتعلقة به . في هذه الفقرة سنستعرض مفهوم آخر يستند بالاساس الى التوزيع المشترك وهو « التوزيع الحدي » .

افرض ان X_1, X_2, X_3 متغيرات عشوائية تسلك وفق دالة توزيع مشترك مثل $P(x_1, x_2, x_3)$ او $f(x_1, x_2, x_3)$. وتطلب الامر ايجاد التوزيع الاحتمالي لاي متغير منها او لاي متغيرين منها . عندئذٍ فان الدالة $P(x_i)$ او $f(x_i)$, $i = 1, 2, 3$. تسمى التوزيع الحدي (الدالة الحدية) للمتغير X_i . كذلك فان الدالة $P(x_1, x_2)$ أو $f(x_1, x_2)$ مثلاً تسمى التوزيع المشترك الحدي

للمتغيرين X_1 , X_2 . ووفق نفس المفهوم الموضح اعلاه يمكن التعميم لحالة K من المتغيرات العشوائية . ان عملية ايجاد دالة التوزيع الحدي تعني انتزاع دالة احتمالية لاي متغير او مجموعة من المتغيرات وهذه العملية تتم على النحو التالي وبفرض وجود توزيع مشترك بثلاث متغيرات . ففي حالة المتغيرات المتقطعة فان :

$$P(x_1) = \sum_{x_2} \sum_{x_3} P(x_1, x_2, x_3), P(x_2) = \sum_{x_1} \sum_{x_3} P(x_1, x_2, x_3)$$

كذلك فان

$$P(x_1, x_2) = \sum_{x_3} P(x_1, x_2, x_3), P(x_2, x_3) = \sum_{x_1} P(x_1, x_2, x_3)$$

اما في حالة المتغيرات المستمرة فان

$$f(x_1) = \int_{x_2} \int_{x_3} f(x_1, x_2, x_3)\, dx_3 \, . \, dx_2$$

وان

$$f(x_3) = \int_{x_1} \int_{x_2} f(x_1, x_2, x_3)\, dx_2 dx_1$$

كذلك فان

$$f(x_1, x_2) = \int_{x_3} f(x_1, x_2, x_3)\, dx_3$$

وان

$$f(x_2, x_3) = \int_{x_1} f(x_1, x_2, x_3)\, dx_1$$

ويمكن تعميم الصيغ اعلاه في حالة وجود توزيع مشترك تتضمن دالته اكثر من ثلاث متغيرات . ان الدوال الحدية المستنتجة وفق الصيغ اعلاه أو غيرها هي ايضاً دوال احتمالية تتصف بالخصائص الثلاثة من حيث كونها دوال وحيدة القيمة . غير سالبة Non – negative function المجموع او التكامل حول قيم متغيرات الدالة

الحدية الممكنة يجب ان يكون مساوياً للواحد . كذلك فان الدوال الحدية ذات **متغير واحد هي في الحقيقة نفس الدوال الاحتمالية التي سبق لنا دراستها في الفقرة** (١ ـ ٤) . في حين ان الدوال الحدية ذات متغيرين او اكثر هي نفس الدوال الاحتمالية المشتركة التي سبق لنا دراستها في الفقرة (٤ ـ ١) . عليه وفي حالة تطلب الامر حساب عزوم للدوال الحدية (أو أي شيء آخر يتعلق بالتوزيع الحدي) فانه يتم الاستعانة بهاتين الفقرتين .

مثال (١١) : افرض أن

$$P(x_1, x_2, x_3) = \frac{1}{36}(x_1 + 2x_2 - x_3);$$

$$x_1 = 1, 2; \; x_2 = 0, 1, 2, \; x_3 = 0, 1$$

تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرات X_1, X_2, X_3 . جد الدالة الحدية الى X_2 ، X_3 ، X_2, X_3 .

الحل

$$P(x_2) = \sum_{x_1=1}^{2} \sum_{x_3=0}^{1} \frac{1}{36}(x_1 + 2x_2 - x_3)$$

$$= \frac{1}{36}\left(\sum_{x_1=1}^{2} \sum_{x_3=0}^{1} x_1 + 2 \sum_{x_1=1}^{2} \sum_{x_3=0}^{1} x_2 - \sum_{x_1=1}^{2} \sum_{x_3=0}^{1} x_3 \right)$$

$$= \frac{1}{36}(6 + 8x_2 - 2) = \frac{1}{9}(2x_2 + 1) \; ; \; x_2 = 0, 1, 2 .$$

لاحظ ان

$$\sum_{x_2=0}^{2} \frac{1}{9}(2x_2 + 1) = \frac{1}{9}(1 + 3 + 5) = 1$$

كذلك

$$P(x_2, x_3) = \frac{1}{36} \sum_{x_1=1}^{2} (x_1 + 2x_2 - x_3)$$

$$= \frac{1}{36}\left(\sum_{x_1=1}^{2} x_1 + 2 \sum_{x_1=1}^{2} x_2 - \sum_{x_1=1}^{2} x_3 \right) = \frac{1}{36}(3 + 4x_2 - 2x_3)$$

فاذن

$$P(x_2, x_3) = \frac{1}{36}(4x_2 - 2x_3 + 3) ; x_2 = 0, 1, 2 ; x_3 = 0, 1$$

كذلك فان

$$P(x_3) = \sum_{x_1=1}^{2} \sum_{x_2=0}^{2} \frac{1}{36}(x_1 + 2x_2 - x_3)$$

$$= \frac{1}{36}(9 + 12 - 6x_3) = \frac{1}{12}(7 - 2x_3) \; ; \; x_3 = 0, 1$$

مثال (١٢) : ــ افرض ان $f(x_1, x_2, x_3) = 8x_1x_2x_3 \; ; 0 < x_1, x_2, x_3 < 1$
تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرات X_1, X_2, X_3 . جد $f(x_3)$

الحل :

$$f(x_3) = \int_0^1 \int_0^1 8x_1x_2x_3 dx_2 \, dx_1$$

$$= 8x_3 \cdot \int_0^1 x_1 \, dx_1 \cdot \int_0^1 x_2 \, dx_2 = 2x_3 \; ; \; 0 < x_3 < 1$$

ويمكن ملاحظة ان $f(x_1) = 2x_1 , f(x_2) = 2x_2$.
كذلك فان

$$f(x_1, x_2) = \int_0^1 8x_1x_2x_3 \, dx_3 = 8x_1x_2 \int_0^1 x_3 \, dx_3$$
$$= 4x_1x_2 \qquad ; \qquad 0 < x_1, x_2 < 1$$

مثال (١٣) : ــ لتكن $f(x_1, x_2) = 2 : 0 < x_1 < x_2 < 1$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1, X_2. جد الدالة الحدية لكل من X_1, X_2

الحل :

$$f(x_1) = \int_{x_1}^{1} 2dx_2 = 2(1 - x_1) \; ; \; 0 < x_1 < 1$$

وان

$$f(x_2) = \int_{0}^{x_2} 2dx_1 = 2x_2 \; ; \; 0 < x_2 < 1$$

مثال (١٤) : افرض ان دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1 , X_2 موصوفة بالآتي :

(x_1 , x_2) :	(0 , 1)	(0 , 2)	(0 , 3)	(1 , 1)	(1 , 2)	(1 , 3)
$P(x_1 , x_2)$:	0·1	0·2	0·3	0·2	0·1	0·1

جد التوزيع الحدي لكل متغير ثم جد معامل الارتباط البسيط بين X_1 , X_2

الحل : للسهولة نعمل الجدول التالي :

x_1 / x_2	0	1	Σ
1	0·1	0·2	0·3
2	0·2	0·1	0·3
3	0·3	0·1	0·4
Σ	0·6	0·4	1

من هذا الجدول نستنتج ان :

$$P(x_1) = \sum_{x_2=1}^{3} P(x_1, x_2) = 0{\cdot}6 \;\; ; x_1 = 0$$
$$= 0{\cdot}4 \;\; ; x_1 = 1$$

كذلك فان

$$P(x_2) = \sum_{x_1=0}^{1} P(x_1, x_2) = 0{\cdot}3 \;\; ; x_2 = 1$$
$$= 0{\cdot}3 \;\; ; x_2 = 2$$
$$= 0{\cdot}4 \;\; ; x_2 = 3$$

وبهدف حساب معامل الارتباط البسيط فان ذلك يتطلب حساب الوسط والتباين لكل متغير وكذلك العزم المشترك ذو المرتبة الثانية وكما يلي :

$$EX_1 = \sum_{x_1=0}^{1} x_1 P(x_1) = (0)(0{\cdot}6) + (1)(0{\cdot}4) = 0{\cdot}4$$

$$EX_1^2 = \sum_{x_1=0}^{1} x_1^2 P(x_1) = (0)^2(0{\cdot}6) + (1)^2(0{\cdot}4) = 0{\cdot}4$$

فاذن

$$\sigma_1^2 = 0{\cdot}4 - (0{\cdot}4)^2 = 0{\cdot}24$$

كذلك

$$EX_2 = \sum_{x_2=1}^{3} x_2 P(x_2) = (1)(0{\cdot}3) + (2)(0{\cdot}3) + 3(0{\cdot}4) = 2{\cdot}1$$

$$EX_2^2 = \sum_{x_2=1}^{3} x_2^2 P(x_2) = (1)^2(0{\cdot}3) + (2)^2(0{\cdot}3) + (3)^2(0{\cdot}4) = 5{\cdot}1$$

فاذن

$$\sigma_2^2 = 5{\cdot}1 - (2{\cdot}1)^2 = 0{\cdot}69$$

وان

$$EX_1X_2 = (0)(1)(0{\cdot}1) + (0)(2)(0{\cdot}2) + \dots + (1)(3)(0{\cdot}1) = 0{\cdot}7$$

فاذن

$$\sigma_{12} = 0.7 - (0.4)(2.1) = -0.14$$

عليه فان

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 . \sigma_2} = -0.344$$

٤ ـ ٣ : التوزيع الشرْطي Conditional distribution

ان مفهوم التوزيع الشرْطي يقترن بمفهوم الاحتمال الشرْطي المنوه عنه في الفقرة (١ ـ ٢) . حيث انه لاي حادثتين مثل A, B فان احتمال وقوع A علماً ان B ستقع هو $P_r(A \mid B) = P_r(A \cap B) \mid P_r(B)$ وببساطة فان الاحتمال الشرْطي ماهو الا حاصل قسمة « الاحتمال المشترك » على « الاحتمال الحدي » . ووفق هذا المنظور سيتم تعريف التوزيع الاحتمالي الشرْطي : بفرض ان X_1, X_2, X_3 ثلاث متغيرات عشوائية تسلك وفق دالة كتلة او كثافة احتمالية مشتركة وتطلب الامر ايجاد التوزيع الاحتمالي لمجموعة جزئية من هذه المتغيرات بفرض ان المتغيرات الاخرى المتمثلة بمجموعة جزئية متممة للاولى تمتلك قيماً معطاة سلفاً عندئذ فان التوزيع الاحتمالي لهذه المجموعة الجزئية ماهو الا حاصل قسمة دالة التوزيع الاحتمالي المشترك على دالة التوزيع الحدي لتلك المتغيرات المتمثلة بالمجموعة الجزئية المتممة . هذا التوزيع يسمى « التوزيع الشرْطي » . فمثلاً لو تطلب الامر ايجاد دالة التوزيع الشرْطي للمتغير X_1 علماً ان $X_2 = x_2 , X_3 = x_3$ فان ذلك يتم وفق الآتي وبفرض ان المتغيرات من النوع المتقطع :

$$P(x_1 \mid X_2 = x_2 , X_3 = x_3) = \frac{P(x_1, x_2, x_3)}{P(x_2, x_3)}$$

واذا كانت المتغيرات من النوع المستمر فان

$$f(x_1 \mid X_2 = x_2 , X_3 = x_3) = \frac{f(x_1, x_2, x_3)}{f(x_2, x_3)}$$

كذلك فان

$$P(x_1, x_2 \mid X_3 = x_3) = \frac{P(x_1, x_2, x_3)}{P(x_3)}$$

$$f(x_1, x_2 \mid X_3 = x_3) = \frac{f(x_1, x_2, x_3)}{f(x_3)}$$

ويمكن تعميم الصيغ اعلاه لحالة اكثر من ثلاث متغيرات . وبشكل خاص اذا كان X_1 , X_2 متغيرين عشوائيين بدالة احتمالية مشتركة معينة ، فان :

$$P(x_1 \mid X_2 = x_2) = \frac{P(x_1, x_2)}{P(x_2)} , P(x_2 \mid X_1 = x_1) = \frac{P(x_1, x_2)}{P(x_1)}$$

$$f(x_1 \mid X_2 = x_2) = \frac{f(x_1, x_2)}{f(x_2)} , f(x_2 \mid X_1 = x_1) = \frac{f(x_1, x_2)}{f(x_1)}$$

ان التوزيعات الشرطية هي الاخرى تتصف بخواص دوال الكتلة او الكثافة الاحتمالية من حيث كونها دوال وحيدة القيمة . غير سالبة . المجموع او التكامل حول قيم تلك المتغيرات يجب ان يكون مساوياً للواحد . فمثلاً وبفرض ان $f(x_1, x_2)$ تمثل دالة الكثافة الاحتمالية للمتغيرين X_1, X_2 وان التوزيع الشرطي الى X_1 علماً ان $X_2 = x_2$ هو $f(x_1 \mid x_2) = \frac{f(x_1, x_2)}{f(x_2)}$ فان

$$\int_{-\infty}^{\infty} f(x_1 \mid x_2)\, dx_1 = \frac{1}{f(x_2)} \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1$$

$$= \frac{1}{f(x_2)} \cdot f(x_2) = 1$$

مثال (١٥): لمعطيات المثال (١١) فان

$$P(x_1, x_3 \mid X_2 = x_2) = \frac{P(x_1, x_2, x_3)}{P(x_2)} = \frac{\frac{1}{36}(x_1 + 2x_2 - x_3)}{\frac{1}{9}(2x_2 + 1)}$$

$$= \frac{x_1 + 2x_2 - x_3}{4(2x_2 + 1)} ; x_1 = 1,2 ; x_2 = 0,1,2 ; x_3 = 0,1$$

وهذا يعني ان

$$P(x_1, x_3 \mid X_2 = x_2) = \frac{1}{4}(x_1 - x_3)\,;\, x_2 = 0$$

$$= \frac{1}{12}(x_1 - x_3 + 2)\,;\, x_2 = 1$$

$$= \frac{1}{20}(x_1 - x_3 + 4)\,;\, x_2 = 2$$

كذلك فان

$$P(x_1 \mid X_2 = x_2, X_3 = x_3) = \frac{\frac{1}{36}(x_1 + 2x_2 - x_3)}{\frac{1}{36}(3 + 4x_2 - 2x_3)} = \frac{x_1 + 2x_2 - x_3}{3 + 4x_2 - 2x_3}$$

وهذا يعني ان :

$$P(x_1 \mid X_2 = x_2, X_3 = x_3) = \frac{1}{3}x_1\,;\, x_2 = 0, x_3 = 0$$

$$= x_1 - 1\,;\, x_2 = 0, x_3 = 1$$

$$= \frac{1}{7}(x_1 + 2)\,;\, x_2 = 1, x_3 = 0$$

$$= \frac{1}{5}(x_1 + 1)\,;\, x_2 = 1, x_3 = 1$$

$$= \frac{1}{11}(x_1 + 4)\,;\, x_2 = 2, x_3 = 0$$

$$= \frac{1}{9}(x_1 + 3)\,;\, x_2 = 2, x_3 = 1$$

ويطلب من القاريء البيان ان $P(x_3 \mid X_1 = x_1, X_2 = x_2) = \frac{x_1 + 2x_2 - x_3}{2x_1 + 4x_2 - 1}$

مثال (١٦) : لمعطيات المثال (١٣) فان :

$$f(x_1 \mid X_2 = x_2) = \frac{2}{2x_2} = \frac{1}{x_2} \; ; 0 < x_1 < x_2 \, , 0 < x_2 < 1$$

وان

$$f(x_2 \mid X_1 = x_1) = \frac{2}{2(1 - x_1)} = \frac{1}{1 - x_1} \; ; x_1 < x_2 < 1 , 0 < x_1 < 1$$

كذلك فان

$$f(x_1 \mid X_2 = x_2) = 4 \quad ; \; x_2 = \frac{1}{4} \, , 0 < x_1 < \frac{1}{4}$$

$$= 2 \quad ; \; x_2 = \frac{1}{2} \, , 0 < x_1 < \frac{1}{2}$$

$$= \frac{4}{3} \quad ; \; x_2 = \frac{3}{4} \, , 0 < x_1 < \frac{3}{4}$$

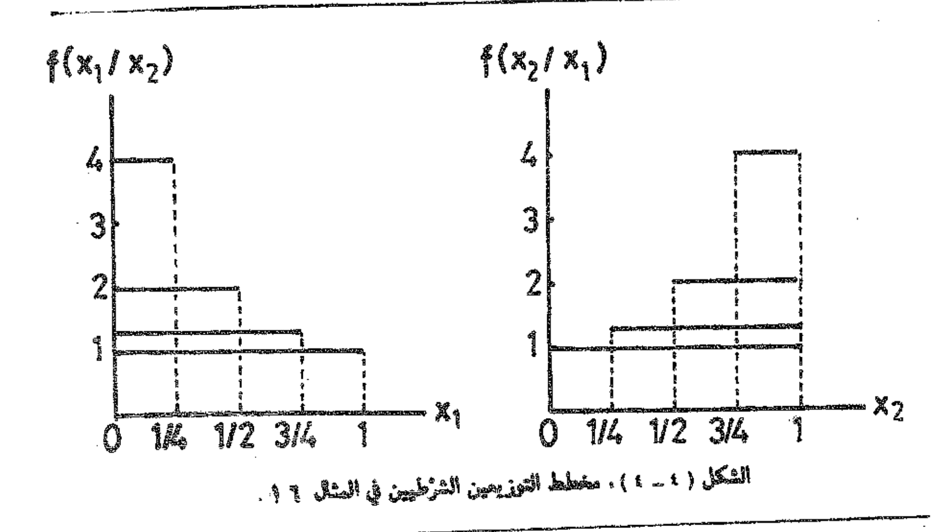
وان

$$f(x_2 \mid X_1 = x_1) = 1 \quad ; \; x_1 = 0 \quad , 0 < x_2 < 1$$

$$= \frac{4}{3} \quad ; \; x_1 = \frac{1}{4} \, , \; \frac{1}{4} < x_2 < 1$$

$$= 2 \quad ; \; x_1 = \frac{1}{2} \, , \; \frac{1}{2} < x_2 < 1$$

والشكل (٤ ـ ٤) يوضح مخطط هاتين الدالتين :



الشكل (٤ ـ ٤) ، مخطط التوزيعين الشرطيين في المثال ١٦ .

٤ ـ ٣ ـ ١ : الاحتمال الشرْطي Conditional probability

لتكن X_1, X_2, X_3 ثلاث متغيرات عشوائية وان $f(x_1, x_2 \mid x_3)$ تمثل دالة الكثافة الاحتمالية المشتركة الشرْطية للمتغيرين X_1, X_2 علماً ان $X_3 = x_3$ وافرض ان $[a_1, b_1]$ مجموعة جزئية معرفة على فضاء X_1 وان $[a_2, b_2]$ مجموعة جزئية معرفة على فضاء X_2 . لتكن A , B حادثتين معرفتين بالشكل $A=\{x_1 : a_1 \leq x_1 \leq b_1\}$ $B=\{x_2: a_2 \leq x_2 \leq b_2\}$ عندئذ فان احتمال وقوع A , B معاً علماً ان $X_3 = x_3$ يسمى « الاحتمال المشترك الشرْطي » للحادثتين A, B ، اي $P_r(A \cap B \mid X_3 = x_3)$.

ويتم حساب هذا الاحتمال وفق ما يلي

$$P_r(A \cap B \mid X_3 = x_3) = P_r(a_1 \leq X_1 \leq b_1 , a_2 \leq X_2 \leq b_2 \mid X_3 = x_3)$$

اذا كانت المتغيرات من النوع المتقطع

$$= \sum_{x_1=a_1}^{b_1} \sum_{x_2=a_2}^{b_2} P(x_1, x_2 \mid X_3 = x_3)$$

اذا كانت المتغيرات من النوع المستمر

$$= \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1, x_2 \mid X_3 = x_3)\, dx_2 dx_1$$

ان المفهوم اعلاه يمكن تطبيقه على اية حالة اخرى تتضمن اكثر من ثلاثة متغيرات . وبشكل خاص اذا كان X_2 , X_1 متغيرين عشوائيين فان :

$$P_r(a_1 \le X_1 \le b_1 \mid X_2 = x_2) = \sum_{x_1=a_1}^{b_1} P(x_1 \mid X_2 = x_2)$$ لمتغيرات متقطعة

$$= \int_{a_1}^{b_1} f(x_1 \mid X_2 = x_2)\, dx_2$$ لمتغيرات مستمرة

مثال (١٧) : لمعطيات المثال (١٥) جد ما يلي :

$$P_r(X_1 = 1, X_3 = 1 \mid X_2 = x_2),\ P_r(X_1 = 1 \mid X_2 = x_2, X_3 = x_3)$$

الحل :

$$P_r(X_1 = 1, X_3 = 1 \mid X_2 = x_2) = \left.\frac{x_1 + 2x_2 - x_3}{4(2x_2 + 1)}\right]_{x_1 = 1, x_3 = 1}$$

$$= \frac{1}{2(2x_2 + 1)}\ ; x_2 = 0, 1, 2$$

$$= \frac{1}{2}\ ; x_2 = 0,\ = \frac{1}{6}\ ; x_2 = 1,\ = \frac{1}{10}\ ; x_2 = 2$$

كذلك فان

$$P_r(X_1 = 1 \mid X_2 = x_2, X_3 = x_3) = \left.\frac{x_1 + 2x_2 - x_3}{3 + 4x_2 - 2x_3}\right]_{x_1 = 1}$$

$$= \frac{2x_2 - x_3 + 1}{4x_2 - 2x_3 + 3}\ ; x_2 = 0, 1, 2\ ; x_3 = 0, 1$$

$$= P$$

لاحظ ان

$(x_2, x_3) : (0,0), (0,1), (1,0), (1,1), (2,0), (2,1)$

$P : \frac{1}{3}, 0, \frac{3}{7}, \frac{2}{5}, \frac{5}{11}, \frac{4}{9}$

مثال (١٨) : لمعطيات المثال (١٦) فان :

$$P_r(a < X_1 < b \mid X_2 = x_2) = \int_a^b \frac{1}{x_2} dx_1 = \frac{b-a}{x_2} \; ; x_2 \geq b - a$$

فمثلاً

$$P_r(0{\cdot}1 < X_1 < 0{\cdot}8 \mid X_2 = x_2) = \frac{0{\cdot}7}{x_2} \quad ; x_2 \geq 0{\cdot}7$$

$$= 0{\cdot}875 \quad ; x_2 = 0{\cdot}8$$

$$= 0{\cdot}778 \quad ; x_2 = 0{\cdot}9$$

وان

$$P_r(0 < X_1 < 0{\cdot}3 \mid X_2 = x_2) = \frac{0{\cdot}3}{x_2} \quad ; x_2 \geq 0{\cdot}3$$

$$= 0{\cdot}6 \quad ; x_2 = 0{\cdot}5$$

$$= 0{\cdot}375 \quad ; x_2 = 0{\cdot}8$$

كذلك فان

$$P_r(c < X_2 < d \mid X_1 = x_1) \int_c^d \frac{1}{1-x_1} dx_2 = \frac{d-c}{1-x_1} \; ; x_1 \leq 1 - (d - c)$$

فمثلاً

$$P_r(0{\cdot}3 < X_2 < 0{\cdot}7 \mid X_1 = x_1) = \frac{0{\cdot}4}{1 - x_1} \quad ; x_1 \leq 0{\cdot}6$$

$$= 0{\cdot}8 \quad ; x_1 = 0{\cdot}5$$

$$= 0{\cdot}5 \quad ; x_1 = 0{\cdot}2$$

وان

$$P_r(0{\cdot}5 < X_2 < 1 \mid X_1 = x_1) = \frac{0{\cdot}5}{1 - x_1} \quad ; x_1 \leq 0{\cdot}5$$

$$= 0{\cdot}833 \quad ; x_1 = 0{\cdot}4$$

$$= 0{\cdot}556 \quad ; x_1 = 0{\cdot}1$$

٤ ـ ٣ ـ ٢ : الدالة التوزيعية الشرْطية
Conditional distribution function

لتكن X_1, X_2, X_3 ثلاثة متغيرات عشوائية وان $f(x_1, x_2 | X_3 = x_3)$ تمثل دالة التوزيع المشترك الشرْطي للمتغيرين X_1, X_2 علماً ان $X_3 = x_3$. عندئذ تعرف الدالة التوزيعية المشتركة الشرْطية للمتغيرين X_1, X_2 على النحو التالي

$$F(x_1, x_2 | X_3 = x_3) = P_r(X_1 \leq x_1, X_2 \leq x_2 | X_3 = x_3)$$

$$= \sum_{-\infty}^{x_1} \sum_{-\infty}^{x_2} P(u_1, u_2 | X_3 = x_3)$$ في حالة المتغيرات المتقطعة

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(u_1, u_2 | X_3 = x_3)\, du_2\, du_1$$ في حالة المتغيرات المستمرة

ونفس هذا المفهوم ينطبق على حالة وجود اكثر من ثلاث متغيرات . وبشكل خاص اذا كان X_1, X_2 متغيرين عشوائيين فان :

$$F(x_1 | X_2 = x_2) = \sum_{-\infty}^{x_1} p(u_1 | X_2 = x_2)$$ لمتغيرات متقطعة

$$= \int_{-\infty}^{x_1} f(u_1 | X_2 = x_2)\, du_1$$ لمتغيرات مستمرة

واذا كانت المتغيرات من النوع المستمر فان

$$\frac{\partial^2 F(x_1, x_2 | X_3 = X_3)}{\partial x_1 . \partial x_2} = f(x_1, x_2 | X_3 = x_3)$$

وان

$$\frac{\partial F(x_1 | X_2 = x_2)}{\partial x_1} = f(x_1 | X_2 = x_2)$$

علماً ان الخصائص المنوه عنها في الفقرة (١ ـ ٥) و (٤ ـ ١ ـ ٣) تتحقق جميعاً هنا وبمجرد اعادة الترميز الى $F(\ldots|\ldots)$ لذا ارتأينا عدم ذكرها تجنباً للتكرار .

مثال (١٩) : لمعطيات المثال (١٦) فان :

$$F(x_1 \mid X_2 = x_2) = P_r(X_1 \leq x_1 \mid X_2 = x_2) = \int_0^{x_1} \frac{1}{x_2}\, du_1$$

$$= \frac{x_1}{x_2} \; ; 0 < x_1 < x_2 \; ; 0 < x_2 < 1$$

لاحظ ان :

$$F(x_2 \mid X_2 = x_2) = \left. \frac{x_1}{x_2} \right]_{x_1 = x_2} = 1$$

وان

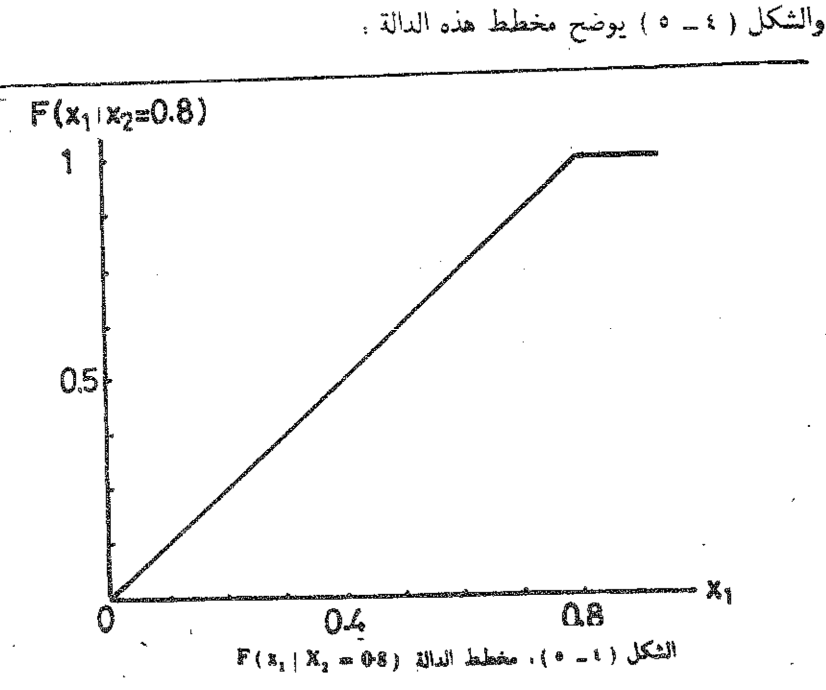
$$F(0 \mid X_2 = x_2) = \left. \frac{x_1}{x_2} \right]_{x_1 = 0} = 0$$

$$\frac{\partial F(x_1 \mid X_2 = x_2)}{\partial x_1} = \frac{1}{x_2}$$

واذا كانت $X_2 = 0.8$ فان

$$F(x_1 \mid X_2 = 0.8) = \frac{x_1}{0.8} \;, x_1 \leq 0.8$$

والشكل (٤ ـ ٥) يوضح مخطط هذه الدالة :



الشكل (٤ ـ ٥) ، مخطط الدالة $F(x_1 \mid X_2 = 0.8)$

كذلك فان

$$F(x_2 \mid X_1 = x_1) = P_r(X_2 \leq x_2 \mid X_1 = x_1)$$

$$= \int_{x_1}^{x_2} \frac{1}{1 - x_1}\, du_2 = \frac{x_2 - x_1}{1 - x_1} \;;\; x_1 < x_2 < 1 \,,\, 0 < x_1 < 1$$

لاحظ هنا ان

$$F(x_1 \mid X_1 = x_1) = \left. \frac{x_2 - x_1}{1 - x_1} \right]_{x_2 = x_1} = 0$$

وان

$$F(1 \mid X_1 = x_1) = \left. \frac{x_2 - x_1}{1 - x_1} \right]_{x_2 = 1} = 1$$

كذلك فان

$$\frac{\partial F(x_2 \mid X_1 = x_1)}{\partial x_2} = \frac{1}{1 - x_1}$$

واذا كانت $x_1 = 0{\cdot}2$ فان

$$F(x_2 \mid X_1 = 0{\cdot}2) = \frac{x_2 - 0{\cdot}2}{0{\cdot}8} \;;\; x_2 \geq 0{\cdot}2$$

والشكل (٤ ـ ٦) يوضح مخطط هذه الدالة

$F(x_2 \mid x_1 = 0.2)$

1

0.5

0 0.2 1 x_2

الشكل (٤ ـ ٦)، مخطط الدالة $F(x_2 \mid X_1 = 0{\cdot}2) = \frac{x_2 - 0{\cdot}2}{8{\cdot}0}$

٤ ـ ٣ ـ ٣ : التوقع الشرْطي وتطبيقاته Conditional expectation

ليكن X_1 , X_2 متغيرين عشوائيين بدالة كتلة احتمالية شرْطية $P(x_1 | X_2 = x_2)$. او كثافة احتمالية شرطية مثل $f(x_1 | X_2 = x_2)$ وان $g(x_1)$ دالة معينة بدلالة المتغير X_1 . عندئذ يعرف التوقع الشرْطي للدالة g بانه « متوسط » هذه الدالة المشروط بقيمة $X_2 = x_2$ ويتم حساب هذا التوقع وفق الآتي :

$$E[g(x_1)|X_2 = x_2] = \sum_{x_1} g(x_1) \cdot p(x_1 | X_2 = x_2) \quad \text{لمتغيرات متقطعة}$$

$$= \int_{-\infty}^{\infty} g(x_1) f(x_1 | X_2 = x_2)\, dx_1 \quad \text{لمتغيرات مستمرة}$$

علماً انه يمكن تعميم التعريف اعلاه لحالة وجود اكثر من متغيرين . وبشكل خاص اذا كانت $g(x_1) = x_1$ فان $E(X_1 | X_2 = x_2)$ يسمى المتوسط الشرْطي Conditional mean لقيم X_1 المشروط بقيمة $X_2 = x_2$. وبالمثل فان $E(X_2 | X_1 = x_1)$ يمثل المتوسط الى X_2 المشروط بقيمة $X_1 = x_1$. وهذا يعني ان

$$E(X_1 | X_2 = x_2) = \sum_{x_1} x_1\, p(x_1 | X_2 = x_2) \quad \text{لمتغيرات متقطعة}$$

$$= \int_{-\infty}^{\infty} x_1\, f(x_1 | X_2 = x_2)\, dx_1 \quad \text{لمتغيرات مستمرة}$$

واذا كانت $g(x_1) = x_1^2$ فان $E(X_1^2 | X_2 = x_2)$ يسمى « متوسط مربعات » قيم X_1 المشروط بقيمة $X_2 = x_2$. وهذا يعني ان :

$$E(X_1^2 | X_2 = x_2) = \sum_{x_1} x_1^2\, p(x_1 | X_2 = x_2) \quad \text{لمتغيرات متقطعة}$$

$$= \int_{-\infty}^{\infty} x_1^2\, f(x_1 | X_2 = x_2)\, dx_1 \quad \text{لمتغيرات مستمرة}$$

وبذلك يمكن تعريف « التباين الشرْطي conditional variance
استناداً لما تقدم وفق مايلي :

$$\sigma^2_{1\cdot 2} = V(X_1 \mid X_2 = x_2) = E(X_1^2 \mid X_2 = x_2) - [E(X_1 \mid X_2 = x_2)]^2$$

وبشكل عام فان $E(X_1^r \mid X_2 = x_2)$ يسمى « العزم الشرْطي ذا المرتبة r حول نقطة الاصل للمتغير X_1 المشروط بقيمة $X_2 = x_2$
واذا كانت X_1, X_2, X_3 ثلاثة متغيرات عشوائية فان $E(X_1 X_2 \mid X_3 = x_3)$ يسمى « العزم المشترك الشرْطي بين X_1, X_2 المشروط بقيمة $X_3 = x_3$. ويتم حساب هذا العزم وفق الصيغة : ــ
لمتغيرات متقطعه

$$E(X_1 X_2 \mid X_3 = x_3) = \sum_{x_1} \sum_{x_2} x_1 x_2 \, p(x_1, x_2 \mid X_3 = x_3)$$

لمتغيرات مستمرة

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2 \mid X_3 = x_3) \, dx_2 \, dx_1$$

وبشكل عام فان $E(X_1^r X_2^m \mid X_3 = x_3)$ يسمى العزم المشترك الشرْطي ذا المرتبة $(r + m)$ المشروط بقيمة $X_3 = x_3$ وهذا يعني انه يمكن حساب قيمة « التباين المشترك الشرْطي » بين X_1, X_2 المشروط بقيمة $X_3 = x_3$ وفق الصيغة التالية :

$$\sigma_{12\cdot 3} = \text{cov}(X_1, X_2 \mid X_3 = x_3)$$

$$= E(X_1 X_2 \mid X_3 = x_3) - E(X_1 \mid X_3 = x_3) . E(X_2 \mid X_3 = x_3)$$

وعلى ضوء مفهوم التباين المشترك الشرْطي يمكن قياس درجة العلاقة بين X_1, X_2 علماً ان $X_3 = x_3$ وفق مايسمى بمعامل الارتباط الجزئي المنوه عنه في الفقرة (٤ ــ ١ ــ ٥) وذلك من خلال الصيغة :

$$\rho_{12\cdot 3} = \frac{\sigma_{12\cdot 3}}{\sigma_{1\cdot 3} . \sigma_{2\cdot 3}}$$

حيث $\sigma_{1\cdot 3}$ ، $\sigma_{2\cdot 3}$ يمثلان على التوالي الانحراف المعياري الى كل من X_1, X_2 المشروط بقيمة $X_3 = x_3$ اللذين يتم الحصول عليهما من خلال مايلي :

$$\sigma^2_{1\cdot3} = E(X_1^2 \mid X_3 = x_3) - [E(X_1 \mid X_3 = x_3)]^2$$

$$\sigma^2_{2\cdot3} = E(X_2^2 \mid X_3 = x_3) - [E(X_2 \mid X_3 = x_3)]^2$$

علماً انه يمكن تعميم صيغ حساب التباين المشترك الشرطي ومعامل الارتباط الجزئي في حالة وجود اكثر من ثلاثة متغيرات عشوائية.

مثال (٢٠) : اذا علمت ان $p(x_1 \mid x_2, x_3) = \dfrac{x_1 + 2x_2 - x_3}{3 + 4x_2 - 2x_3}$

$x_1 = 1,2\,,\ x_2 = 0, 1, 2\,,\ x_3 = 0,1$ جد $EX_1\,,\ EX_1^2\,,\ V(X_1)$.

$$EX_1 = E(X_1 \mid X_2 = x_2, X_3 = x_3) = \sum_{x_1=1}^{2} x_1 \cdot \frac{x_1 + 2x_2 - x_3}{3 + 4x_2 - 2x_3}$$

$$= \frac{5 + 6x_2 - 3x_3}{3 + 4x_2 - 2x_3}\,,\ x_2 = 0, 1, 2\,,\ x_3 = 0,1$$

واضح ان :

$$\begin{aligned}
EX_1 &= 5/3\,,\ x_2 = 0\,,\ x_3 = 0\\
&= 2\,,\ x_2 = 0\,,\ x_3 = 1\\
&= 11/7\,,\ x_2 = 1\,,\ x_3 = 0\\
&= 8/5\,,\ x_2 = 1\,,\ x_3 = 1\\
&= 17/11\,,\ x_2 = 2\,,\ x_3 = 0\\
&= 14/9\,,\ x_2 = 2\,,\ x_3 = 1
\end{aligned}$$

$$EX_1^2 = E(X_1^2 \mid X_2 = x_2, X_3 = x_3) = \sum_{x_1=1}^{2} x_1^2 \cdot \frac{x_1 + 2x_2 - x_3}{3 + 4x_2 - 2x_3}$$

$$= \frac{9 + 10x_2 - 5x_3}{3 + 4x_2 - 2x_3}\,,\ x_2 = 0, 1, 2,\,,\ x_3 = 0,1$$

واضح ان :

$$\begin{aligned}
EX_1^2 &= 3\,,\ x_2 = 0\,,\ x_3 = 0\\
&= 4\,,\ x_2 = 0\,,\ x_3 = 1
\end{aligned}$$

$$= 19/7 \quad , x_2 = 1 , x_3 = 0$$
$$= 14/5 \quad , x_2 = 1 , x_3 = 1$$
$$= 29/11 \; , x_2 = 2 , x_3 = 0$$
$$= 24/9 \quad , x_2 = 2 , x_3 = 1$$

كذلك فان

$$V(X_1) = V(X_1 | X_2 = x_2, X_3 = x_3)$$

$$= \frac{9 + 10x_2 - 5x_3}{3 + 4x_2 - 2x_3} - \left[\frac{5 + 6x_2 - 3x_3}{3 + 4x_2 - 2x_3} \right]^2$$

واضح ان :

$$V(X_1) = 2/9 \quad , \; x_2 = 0 , x_3 = 0$$
$$= 0 \quad , \; x_2 = 0 , x_3 = 1$$
$$= 12/49 \; , \; x_2 = 1 , x_3 = 0$$
$$= 6/25 \quad , \; x_2 = 1 , x_3 = 1$$
$$= 30/121 , \; x_2 = 2 , x_3 = 0$$
$$= 20/81 \; , \; x_2 = 2 , x_3 = 1$$

مثال (٢١) : اذا علمت ان

$$p(x_1 / x_2) = \frac{2x_1 + 4x_2 - 1}{4(2x_2 + 1)} \; , p(x_3 | x_2) = \frac{4x_2 - 2x_3 + 3}{4(2x_2 + 1)}$$

وان $p(x_1, x_3 | x_2) = \frac{x_1 + 2x_2 - x_3}{4(2x_2 + 1)}, x_1 = 1, 2, x_2 = 0, 1, 2, x_3 = 0,1$

جد $\sigma_{13 \cdot 2}$ ثم احسب $\rho_{13 \cdot 2}$

الحل :

$$E(X_1 X_3 | X_2 = x_2) = \sum_{x_1=1}^{2} \sum_{x_3=0}^{1} x_1 x_3 \cdot \frac{x_1 + 2x_2 - x_3}{4(2x_2 + 1)}$$

$$= \frac{3x_2 + 1}{2(2x_2 + 1)} \; , x_2 = 0, 1, 2$$

$$E(X_1 \mid X_2 = x_2) = \sum_{x_1=1}^{2} x_1 \cdot \frac{2x_1 + 4x_2 - 1}{4(2x_2 + 1)} = \frac{12x_2 + 7}{4(2x_2 + 1)}$$

$, x_2 = 0, 1, 2$

$$E(X_3 \mid X_2 = x_2) = \sum_{x_3=0}^{1} x_3 \cdot \frac{4x_2 - 2x_3 + 3}{4(2x_2 + 1)} = \frac{4x_2 + 1}{4(2x_2 + 1)}$$

$, x_2 = 0, 1, 2$

$$\therefore \sigma_{13\cdot 2} = \frac{3x_2 + 1}{2(2x_2 + 1)} - \left[\frac{12x_2 + 7}{4(2x_2 + 1)}\right]\left[\frac{4x_2 + 1}{4(2x_2 + 1)}\right]$$

$$= \frac{1}{16(2x_2 + 1)^2}$$

واضح ان :

$$\begin{aligned} \sigma_{13\cdot 2} &= 1/16 \quad ; \ x_2 = 0 \\ &= 1/144 \quad ; \ x_2 = 1 \\ &= 1/400 \quad ; \ x_2 = 2 \end{aligned}$$

كذلك فان

$$E(X_1^2 \mid X_2 = x_2) = \sum_{x_1=1}^{2} x_1^2 \cdot \frac{2x_1 + 4x_2 - 1}{4(2x_2 + 1)} = \frac{20x_2 + 13}{4(2x_2 + 1)}$$

$$E(X_3^2 \mid X_2 = x_2) = \sum_{x_3=0}^{1} x_3^2 \cdot \frac{4x_2 - 2x_3 + 3}{4(2x_2 + 1)} = \frac{4x_2 + 1}{4(2x_2 + 1)}$$

$$\therefore \sigma_{1\cdot 2}^2 = E(X_1^2 \mid X_2 = x_2) - [E(X_1 \mid X_2 = x_2)]^2 = \frac{(4x_2 + 3)(4x_2 + 1)}{16(2x_2 + 1)^2}$$

وان

$$\sigma_{3\cdot 2}^2 = E(X_3^2 \mid X_2 = x_2) - [E(X_3 \mid X_2 = x_2)]^2 = \frac{(4x_2 + 3)(4x_2 + 1)}{16(2x_2 + 1)^2}$$

$$\therefore \rho_{13\cdot 2} = \frac{\sigma_{13\cdot 2}}{\sigma_{1\cdot 2} \cdot \sigma_{3\cdot 2}} = \frac{1}{(4x_2 + 3)(4x_2 + 1)} \ ; x_2 = 0, 1, 2$$

واضح ان

$$\rho_{13\cdot 2} = 1/3 \quad ; \ x_2 = 0$$
$$= 1/35 \quad ; \ x_2 = 1$$
$$= 1/99 \quad ; \ x_2 = 2$$

مثال (٢٢) : اذا علمت ان $f(x_1 \mid X_2 = x_2) = \frac{1}{x_2} ; 0 < x_1 < x_2 , 0 < x_2 < 1$

جد الوسط الشرْطي والتباين الشرطي الى X_1.

الحل :

$$E(X_1 \mid X_2 = x_2) = \int_0^{x_2} \frac{x_1}{x_2} dx_1 = \frac{x_2}{2} \ ; \ 0 < x_2 < 1$$

واضح ان

$$E(X_1 \mid X_2 = x_2) = 1/8 \ ; \ x_2 = 1/4$$
$$= 1/4 \ ; \ x_2 = 1/2$$
$$= 3/8 \ ; \ x_2 = 3/4$$

وان

$$E(X_1^2 \mid X_2 = x_2) = \int_0^{x_2} \frac{x_1^2}{x_2} dx_1 = \frac{x_2^2}{3} \ ; \ 0 < x_2 < 1$$

واضح ان

$$E(X_1^2 \mid X_2 = x_2) = 1/48 \ ; \ x_2 = 1/4$$
$$= 1/12 \ ; \ x_2 = 1/2$$
$$= 3/16 \ ; \ x_2 = 3/4$$

فاذن

$$V(X_1 \mid X_2 = x_2) = \frac{x_2^2}{3} - \left(\frac{x_2}{2}\right)^2 = \frac{x_2^2}{12} \ ; \ 0 < x_2 < 1$$

واضح ان

$$V(X_1 \mid X_2 = x_2) = 1/192 \ ; \ x_2 = 1/4$$
$$= 1/48 \ ; \ x_2 = 1/2$$
$$= 3/64 \ ; \ x_2 = 3/4$$

٤ ـ ٣ ـ ٤ : الدالة المولدة لعزوم التوزيع الشرْطي Moment generating function for conditional distribution

ليكن X_1 , X_2 ، متغيرين عشوائيين بدالة كتلة احتمالية شرْطية مثل $P(X_1 | X_2 = x_2)$ ، او دالة كثافة احتمالية شرْطية مثل $f(X_1 | X_2 = x_2)$.
وافرض ان t متغير آخر وان h عدد موجب بحيث ان $-h < t < h$ عندئذٍ تعرف الدالة المولدة لعزوم التوزيع الشرْطي (اذا كانت موجودة) على النحو التالي :

$$M_{X_1|X_2}(t) = E(e^{tX_1} | X_2 = x_2)$$

$$= \sum_{x_1} e^{tx_1} . P(x_1 | X_2 = x_2) \qquad \text{لمتغيرات متقطعة}$$

$$= \int_{-\infty}^{\infty} e^{tx_1} . f(x_1 | X_2 = x_2) \, dx_1 \qquad \text{لمتغيرات مستمرة}$$

لاحظ ان $M_{X_1|X_2}(t = 0) = 1$

واذا كانت $P(x_1 , x_2 | X_3 = x_3)$ تمثل دالة الكتلة الاحتمالية المشتركة الشرْطية للمتغيرين X_1 , X_2 علماً ان $X_3 = x_3$ ، وان $f(x_1 , x_2 | X_3 = x_3)$ تمثل دالة الكثافة الاحتمالية المشتركة الشرْطية للمتغيرين X_1 , X_2 علما ان $X_3 = x_3$ عندئذٍ فان الدالة المولدة للعزوم المشتركة الشرْطية هي :

$$M_{X_1,X_2|X_3}(t_1 , t_2) = E(e^{t_1X_1 + t_2X_2} | X_3 = x_3)$$

$$= \sum_{x_1} \sum_{x_2} e^{t_1x_1 + t_2x_2} P(x_1 , x_2 | X_3 = x_3) \qquad \text{لمتغيرات متقطعة}$$

لمتغيرات مستمرة

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{t_1x_1 + t_2x_2} . f(x_1 , x_2 | X_3 = x_3) \, dx_2 dx_1$$

لاحظ هنا ان :

$$M_{X_1,X_2|X_3}(0,0) = 1 , M_{X_1,X_2|X_3}(t_1 , 0) = E(e^{t_1X_1} | X_3 = x_3)$$
$$= M_{X_1|X_3}(t_1) , M_{X_1,X_2|X_3}(0 , t_2) = M_{X_2|X_3}(t_2)$$

ومن خلال هذا النوع من الدوال يمكن الحصول على عزوم التوزيع الشرْطي حول نقطة الاصل وفق نفس الاسلوب الموضح في الفقرتين (٢ ــ ٢ ــ ١) و (٤ ــ ١ ــ ٦) .

وهذا يعني ان العزم الشرْطي ذو المرتبة r حول نقطة الاصل للمتغير X_1 علماً ان $X_2 = x_2$ هو

$$\frac{d^r M_{X_1|X_2}(t)}{dt^r}\Bigg]_{t=0} = M^{(r)}_{X_1|X_2}(0) = E(X_1^r \mid X_2 = x_2) : r = 1, 2, \ldots$$

وان العزم المشترك الشرْطي ذو المرتبة $(r + m)$ للمتغيرين X_1, X_2 علماً ان $X_3 = x_3$ هو

$$E(X_1^r . X_2^m \mid X_3 = x_3) = \frac{\partial^{r+m} M_{X_1,X_2/X_3}(t_1, t_2)}{\partial t_1^r . \partial t_2^m}\Bigg]_{t_1 = t_2 = 0}$$

مثال (٢٣) : لمعطيات المثال (٢٢) يطلب ايجاد الدالة المولدة لعزوم المتغير X_1 علماً ان $X_2 = x_2$.

الحل : يتضح من هذا المثال ان

$$f(x_1 \mid X_2 = x_2) = \frac{1}{x_2} ; 0 < x_1 < x_2 , 0 < x_2 < 1$$

فاذن

$$M_{X_1|X_2}(t) = \int_0^{x_2} e^{t_1 x_1} . \frac{1}{x_2} dx_1 = \frac{e^{t_1 x_2} - 1}{t_1 x_2} , t_1 > 0$$

يلاحظ للوهلة الاولى وعند التعويض عن t_1 بالصفر نحصل على الشكل غير المحدد $\left(\frac{0}{0}\right)$ وهذا امر يخالف احدى اهم خصائص الدوال المولدة للعزوم . الا انه وباستخدام « قاعدة لوبيتل **"L'Hopital's Rule** عندئذٍ فان :

$$\lim_{t_1 \to 0} \frac{e^{t_1 x_2} - 1}{t_1 x_2} = \lim_{t_1 \to 0} \frac{\frac{d}{dt_1}(e^{t_1 x_2} - 1)}{\frac{d}{dt_1}(t_1 x_2)} = \lim_{t_1 \to 0} \frac{x_2 e^{t_1 x_2}}{x_2} = 1$$

ويطلب من القاريء البيان ان :

$$M'_{X_1|x_2}(0) = \frac{x_2}{2} \; , M''_{X_1|x_2}(0) = \frac{x_2^2}{3} .$$

٤ ـ ٤ : الاستقلال التصادفي Stochastic independence

ان مفهوم الاستقلال التصادفي بين المتغيرات العشوائية ذو أهمية كبيرة في الكثير من التطبيقات الاحصائية وخصوصاً عند استنتاج التوزيع الاحتمالي لمتغير عشوائي يعتمد على متغيرين او اكثر . ان هذا المفهوم يقترن بمفهوم الاستقلال بين الحوادث لدى دراستنا لموضوع الاحتمالات حيث لاحظنا بأنه اذا كانت **A** و **B** حادثتان وان $A \cap B$ تمثل حادثة تقاطعهما عندئذٍ يقال ان **A** مستقلة عن **B** اذا كان $P_r(A \cap B) = P_r(A).P(B)$. ووفق هذا المفهوم سيتم تعريف مفهوم الاستقلال التصادفي بين المتغيرات العشوائية .

بفرض ان X_2 , X_1 متغيران عشوائيان بدالة كثافة احتمالية مشتركة $f(x_1, x_2)$ وان $f(x_2), f(x_1)$ تمثل على التوالي الدالة الحدية الى X_2, X_1 عندئذٍ فان التوزيع الشرطي الى X_1 علماً ان $X_2 = x_2$ هو $f(x_1 | X_2 = x_2)$. وان $f(x_1, x_2) = f(x_1 | X_2 = x_2) . f(x_2)$ الان بفرض ان $f(x_1 | X_2 = x_2)$ لاتعتمد على X_2 عندئذٍ فان التوزيع الحدي الى X_1 هو :

$$f(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 = \int_{-\infty}^{\infty} f(x_1 \mid X_2 = x_2) f(x_2)\, dx_2$$

$$= f(x_1 \mid X_2 = x_2) \int_{-\infty}^{\infty} f(x_2)\, dx_2 = f(x_1 \mid X_2 = x_2)$$

وهذا يعني ان $f(x_1, x_2) = f(x_1) . f(x_2)$. ان المفهوم اعلاه ينطبق على حالة المتغيرات المتقطعة بمجرد استبدال عملية التكامل بعملية الجمع. يلاحظ مما تقدم ان التوزيع الشرطي الى X_1 علماً ان $X_2 = x_2$ « مستقل » عن X_2 وهذا معناه ان $f(X_1 \mid X_2 = x_2)$ تمثل الدالة الحدية الى X_1. وعندئذٍ فان دالة التوزيع المشترك يمكن التعبير عنها من خلال حاصل ضرب الدالة الحدية للمتغير X_1 في الدالة الحدية للمتغير X_2. وفي هذه الحالة يقال ان المتغيرين X_1, X_2 «مستقلان تصادفياً». وفيما يلي التعريف العام للاستقلال التصادفي بين المتغيرات العشوائية :

افرض ان $X_1, X_2, \ldots, X_k$ متغيرات عشوائية بدالة كتلة احتمالية مشتركة $P(x_1, x_2, \ldots, x_k)$ او كثافة احتمالية مشتركة $f(x_1, x_2, \ldots, x_k)$ عندئذٍ يقال ان هذه المتغيرات « مستقلة تصادفياً » اذا وفقط اذا امكن التعبير عن دالة التوزيع المشترك لهذه المتغيرات من خلال حاصل ضرب الدوال الحدية لها . اي ان :

لمتغيرات متقطعة

$$P(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} P(x_i)$$

لمتغيرات مستمرة

$$f(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} f(x_i)$$

واذا لم يتحقق هذا الشرط عندئذٍ يقال ان المتغيرات « معتمدة تصادفياً »
. **Stochastically dependent**

مثال (٢٤) : افرض ان $f(x_1, x_2) = e^{-(x_1 + x_2)}; x_1, x_2 \geq 0$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1, X_2 هل يمكن القول ان X_1, X_2 مستقلان تصادفياً ؟

الحل : ان الدالة الحدية لكل من هذين المتغيرين هي :

$$f(x_1) = \int_0^\infty e^{-(x_1+x_2)}\,dx_2 = e^{-x_1} \;;\; x_1 \geq 0$$

$$f(x_2) = \int_0^\infty e^{-(x_1+x_2)}\,dx_1 = e^{-x_2} \;;\; x_2 \geq 0$$

فاذن

$$f(x_1)\,.\,f(x_2) = e^{-x_1}.\,e^{-x_2} = e^{-(x_1+x_2)} = f(x_1, x_2)$$

وهذا يعني ان X_1 , X_2 مستقلان تصادفياً .

مثال (٢٥) : لتكن $f(x_1, x_2) = x_1 + x_2 ; 0 < x_1, x_2 < 1$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1, X_2 . اختبر فيما اذا كان X_1 مستقل تصادفيا عن X_2

الحل :

$$f(x_1) = \int_0^1 (x_1 + x_2)\,dx_2 = x_1 + \frac{1}{2} \;;\; 0 < x_1 < 1$$

وان

$$f(x_2) = \int_0^1 (x_1 + x_2)\,dx_1 = x_2 + \frac{1}{2} \;;\; 0 < x_2 < 1$$

فاذن

$$f(x_1)\,.\,f(x_2) = \left(x_1 + \frac{1}{2}\right)\left(x_2 + \frac{1}{2}\right) \neq f(x_1, x_2)$$

وهذا يعني ان المتغيران X_1, X_2 معتمدين تصادفياً .

مثال (٢٦) : افرض ان $P(x_1, x_2) = \frac{1}{16} \;;\; x_1, x_2 = 1, 2, 3, 4$ تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 اختبر فيما اذا كان X_1 مستقل تصادفياً عن X_2 .

الحل :

$$P(x_1) = \sum_{x_2=1}^{4} \frac{1}{16} = \frac{1}{4} , \ P(x_2) = \sum_{x_1=1}^{4} \frac{1}{16} = \frac{1}{4}$$

فاذن

$$P(x_1) . P(x_2) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16} = P(x_1, x_2)$$

وهذا يعني أن X_1 مستقل تصادفياً عن X_2

مثال (٢٧) : افرض ان $P(x_1, x_2) = \frac{x_1 + x_2}{21} \ ; x_1 = 1, 2, 3; x_2 = 1, 2$ تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 هل يمكن القول ان X_1 مستقل تصادفياً عن X_2

الحل :

$$P(x_1) = \sum_{x_2=1}^{2} \frac{x_1 + x_2}{21} = \frac{2x_1 + 3}{21} \ ; x_1 = 1, 2, 3$$

وان

$$P(x_2) = \sum_{x_1=1}^{3} \frac{x_1 + x_2}{21} = \frac{x_2 + 2}{7} \ ; x_2 = 1, 2$$

فاذن

$$P(x_1) . P(x_2) = \left(\frac{2x_1 + 3}{21} \right) \cdot \left(\frac{x_2 + 2}{7} \right) \neq P(x_1, x_2)$$

وهذا يعني ان X_1, X_2 معتمدين تصادفيا .

ان الاستقلال التصادفي بين المتغيرات العشوائية من حيث المضمون يقترن بمفهوم الحوادث المستقلة المنوه عنه في الفقرة (١ــ ٢ــ ٦) . وفيما يلي بعض النظريات التي تخص الاستقلال التصادفي بين المتغيرات العشوائية . علماً اننا سوف نورد البراهين لحالة المتغيرات المستمرة وذات البرهان ينطبق على حالة المتغيرات المتقطعة بمجرد استبدال عملية التكامل بعملية الجمع .
بفرض أن $X_1, X_2, ..., X_k$ متغيرات عشوائية « مستقلة » بدالة كثافة احتمالية مشتركة $f(x_1, x_2, ..., x_k)$ عندئذٍ :

مبرهنة (١ ـ ٤) : اذا كانت $F(x_1, x_2, \ldots, x_k)$ تمثل الدالة التوزيعية المشتركة لهذه المتغيرات عندئذ فان

$$F(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} F(x_i)$$

البرهان : ان

$$F(x_1, x_2, \ldots, x_k) = P_r(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_k \leq x_k)$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_k} f(u_1, u_2, \ldots, u_k)\, du_k . du_{k-1} \ldots du_1$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_k} f(u_1) . f(u_2) \ldots f(u_k) . du_k . du_{k-1} \ldots du_1$$

$$= \int_{-\infty}^{x_1} f(u_1)\, du_1 . \int_{-\infty}^{x_2} f(u_2)\, du_2 . \ldots \int_{-\infty}^{x_k} f(u_k)\, du_k$$

$$= F(x_1) . F(x_2) \ldots F(x_k) = \prod_{i=1}^{k} F(x_i)$$

مثال (٢٨): اذا علمت ان $F(x_1) = 1 - e^{-x_1}; x_1, x_2 \geq 0$ ، وان $F(x_2) = 1 - e^{-x_2}$ و X_1 مستقل عن X_2 عندئذٍ

$$F(x_1, x_2) = F(x_1) . F(x_2) = (1 - e^{-x_1})(1 - e^{-x_2})$$

واضح ان

$$f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \, \partial x_2} = \frac{\partial F(x_1)}{\partial x_1} . \frac{\partial F(x_2)}{\partial x_2}$$

$$= e^{-x_1} . e^{-x_2} = e^{-(x_1 + x_2)}$$

مبرهنة (٢ ـ ٤) : اذا كان

$P_r(a_1 < X_1 < b_1, a_2 < X_2 < b_2, \ldots, a_k < X_k < b_k)$

يمثل الاحتمال المشترك للحوادث $A_i = \{a_i < X_i < b_i\}, a_i < b_i ; i = 1,2,\ldots,k$

عندئذ :

$$P_r(A_1 \cap A_2 \cap \ldots \cap A_k) = \prod_{i=1}^{k} P_r(A_i) = \prod_{i=1}^{k} P_r(a_i < X_i < b_i)$$

البرهان :

$$P_r(A_1 \cap A_2 \cap \ldots \cap A_k) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_k}^{b_k} f(x_1, x_2, \ldots, x_k) dx_k . dx_{k-1} \ldots dx_1$$

$$= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_k}^{b_k} f(x_1) . f(x_2) . \ldots . f(x_k) dx_k . dx_{k-1} \ldots dx_1$$

$$= \int_{a_1}^{b_1} f(x_1) dx_1 . \int_{a_2}^{b_2} f(x_2) dx_2 \ldots \int_{a_k}^{b_k} f(x_k) dx_k$$

$$= P_r(A_1) . P_r(A_2) . \ldots . P_r(A_k) = \prod_{i=1}^{k} P_r(a_i < X_i < b_i)$$

مثال (٢٩) : اذا علمت ان $f(x_1, x_2) = 4x_1x_2 ; 0 < x_1, x_2 < 1$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1 , X_2 . اختبر فيما اذا كان X_1 مستقل تصادفياً عن X_2 ثم جد $P_r(0 < X_1 < 0.5, 0.3 < X_2 < 0.8)$.

الحل :

$$f(x_1) = \int_0^1 4x_1x_2 dx_2 = 2x_1 ; 0 < x_1 < 1$$

$$f(x_2) = \int_0^1 4x_1x_2 dx_1 = 2x_2 ; 0 < x_2 < 1$$

فاذن

$$f(x_1).f(x_2) = (2x_1)(2x_2) = 4x_1x_2 = f(x_1, x_2)$$

وذلك يعني ان X_1 مستقل تصادفياً عن X_2 . فاذن

$$P_r(0 < X_1 < 0{\cdot}5, 0{\cdot}3 < X_2 < 0{\cdot}8) = P_r(0 < X_1 < 0{\cdot}5).P_r(0{\cdot}3 < X_2 < 0{\cdot}8)$$

لكن

$$P_r(0 < X_1 < 0{\cdot}5) = \int_0^{0\cdot5} 2x_1 dx_1 = 0{\cdot}25$$

وان

$$P_r(0{\cdot}3 < X_2 < 0{\cdot}8) = \int_{0\cdot3}^{0\cdot8} 2x_2 dx_2 = 0{\cdot}55$$

فاذن

$$P_r(0 < X_1 < 0{\cdot}5, 0{\cdot}3 < X_2 < 0{\cdot}8) = (0{\cdot}25)(0{\cdot}55) = 0{\cdot}1375$$

مبرهنة (٣ ــ ٤) : لتكن $g_i(x_i), i = 1, 2, \ldots, k$ دوال بدلالة المتغيرات العشوائية $X_1, X_2, \ldots X_k$. عندئذٍ

$$E \prod_{i=1}^{k} g_i(x_i) = \prod_{i=1}^{k} Eg_i(x_i)$$

$$E \prod_{i=1}^{k} g_i(x_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g_1(x_1).g_2(x_2)\ldots g_k(x_k).f(x_1, x_2, \ldots, x_k)$$

$$dx_k . dx_{k-1} . \ldots . dx_1$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g_1(x_1).g_2(x_2)\ldots g_k(x_k).f(x_1).f(x_2)\ldots f(x_k).dx_k \ldots dx_1$$

$$= \int_{-\infty}^{\infty} g_1(x_1)f(x_1)dx_1 . \int_{-\infty}^{\infty} g_2(x_2)f(x_2)dx_2 \ldots \int_{-\infty}^{\infty} g_k(x_k)f(x_k)dx_k$$

$$= Eg_1(x_1).Eg_2(x_2)\ldots Eg_k(x_k) = \prod_{i=1}^{k} Eg_i(x_i)$$

وبشكل خاص اذا كانت $g_i(x_i) = x_i, i = 1, 2, \ldots, k$ فان

$E \prod_{i=1}^{k} X_i = \prod_{i=1}^{k} EX_i$ وهذا يعني ان

$$COV(X_i, X_j) = EX_iX_j - EX_i . EX_j = EX_i . EX_j - EX_i . EX_j = 0$$

وعندئذٍ فان

$$V(X_i \pm X_j) = V(X_i) + V(X_j)$$

وحيث إن التباين المشترك في هذه الحالة مساوٍ للصفر فذلك يعني ان مصفوفة الارتباطات R المشار اليها في الفقرة (٤ ــ ١ ــ ٥) سوف تأخذ شكل مصفوفة احادية identity matrix . اي ان

$$R = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

عليه يمكن القول انه اذا كان X_i مستقل تصادفياً عن X_j فذلك يعني ن $\sigma_{ij} = 0$. لكن العكس غير صحيح اي مانعنيه انه اذا كان $\sigma_{ij} = 0$ فذلك لايعني ان المتغيرين مستقلان تصادفياً وانما هما متغيران غير مرتبطين uncorrelated. والمثال الاتي يوضح حالة من هذا النوع .

مثال (٣٠) : افرض ان دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 موصوفة بالجدول الآتي :

x_2 \ x_1	-1	0	1	Σ
-2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{3}{16}$
-1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{5}{16}$
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{5}{16}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{3}{16}$
Σ	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	1

واضح من هذا الجدول ان

$$P(x_1) = \frac{6}{16} \; ; \; x_1 = -1 \qquad P(x_2) = \frac{3}{16} \; ; \; x_2 = -2$$

$$= \frac{4}{16} \; ; \; x_1 = 0 \qquad = \frac{5}{16} \; ; \; x_2 = -1$$

$$= \frac{6}{16} \; ; \; x_1 = 1 \qquad = \frac{5}{16} \; ; \; x_2 = 1$$

$$= \frac{3}{16} \; ; \; x_2 = 2$$

وان

$$EX_1X_2 = (-1)(-2)\left(\frac{1}{16}\right) + (0)(-2)\left(\frac{1}{16}\right) + \dots$$

$$+ (1)(2)\left(\frac{1}{16}\right) = 0$$

$$EX_1 = (-1)\left(\frac{6}{16}\right) + (0)\left(\frac{4}{16}\right) + (1)\left(\frac{6}{16}\right) = 0$$

$$EX_2 = (-2)\left(\frac{3}{16}\right) + (-1)\left(\frac{5}{16}\right) + (1)\left(\frac{5}{16}\right)$$

$$+ (2)\left(\frac{3}{16}\right) = 0$$

فاذن

$$\sigma_{12} = EX_1X_2 - EX_1 . EX_2 = 0 \quad \therefore \rho_{12} = 0$$

وهذا يعني ان X_1, X_2 غير مرتبطين . لكن

$$P(x_1 = -1, x_2 = -2) = \frac{1}{16}$$

وان

$$P(x_1 = -1) = \frac{6}{16}, P(x_2 = -2) = \frac{3}{16}$$

فاذن

$$P(x_1 = -1) . P(x_2 = -2) = \frac{6}{16} \cdot \frac{3}{16} = \frac{9}{128}$$

$$\neq P(x_1 = -1, x_2 = -2)$$

$$P(x_1 = -1) . P(x_2 = 1) = \frac{6}{16} \cdot \frac{5}{16} = \frac{12}{128}$$

$$\neq P(x_1 = -1, x_2 = 1) = \frac{1}{8}$$

عليه وبشكل عام فان

$$P(x_1, x_2) \neq P(x_1).P(x_2)$$

وهذا يعني ان X_1, X_2 معتمدان تصادفياً (غير مستقلين) .

مبرهنة (٤ ـ ٤) : اذا كانت $M(t_1, t_2, ..., t_k)$ تمثل الدالة المولدة لعزوم التوزيع المشترك للمتغيرات $X_1, X_2, ..., X_k$ ، عندئذٍ

$$M(t_1, t_2, ..., t_k) = \prod_{i=1}^{k} M_{X_i}(t_i)$$

البرهان :

$$M(t_1, t_2, ..., t_k) = E\, e^{\sum_{i=1}^{k} t_i X_i} = E \prod_{i=1}^{k} e^{t_i X_i}$$

وبفرض ان $g_i(x_i) = e^{t_i x_i}$ وباستخدام المبرهنة (٣ ـ ٤) فذلك يعني ان :

$$M(t_1, t_2, ..., t_k) = \prod_{i=1}^{k} E\, e^{t_i X_i} = \prod_{i=1}^{k} M_{X_i}(t_i)$$

نتيجة لمبرهنة (٤ ـ ٤) :

اذا كانت المتغيرات المستقلة $X_1, X_2, ..., X_k$ تمتلك نفس التوزيع الاحتمالي المعرف بدالة كتلة احتمالية $P(x)$ او دالة كثافة احتمالية $f(x)$ الذي يمتلك دالة مولدة لعزومه حول نقطة الاصل مثل $M_X(t)$ ، عندئذ فان $M_{X_i}(t_i) = M_X(t), \forall_i = 1, 2, ... k$ وان

$$M(t_1, t_2, ..., t_k) = \prod_{i=1}^{k} M_X(t) = [M_X(t)]^k$$

مثال (٣١) : افرض ان $f(x_1, x_2, x_3) = e^{-(x_1+x_2+x_3)}, x_i \geq 0$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرات X_1, X_2, X_3 . جد الدالة المولدة لعزوم هذا التوزيع المشترك .

الحل : يمكن البيان أن هذه المتغيرات مستقلة تصادفياً وان $M_{X_i}(t_i) = \frac{1}{1-t_i}\ ; t_i < 1$ عندئذ واستناداً للنظرية (٤) فان

$$M(t_1, t_2, t_3) = \prod_{i=1}^{3} M_{X_i}(t_i) = \frac{1}{(1-t_1)(1-t_2)(1-t_3)}$$

مثال (٣٢) : افرض ان X_1, X_2, X_3, X_4 تمثل قياسات عينة عشوائية مختارة من توزيع احتمالي بدالة كثافة احتمالية $f(x) = e^{-x}; x \geq 0$
جد الدالة المولدة لعزوم $Y = X_1 + X_2 + X_3 + X_4$

الحل : حيث ان قياسات العينة العشوائية هي بحكم متغيرات عشوائية مستقلة تمتلك نفس التوزيع الاحتمالي فذلك يعني ان

$$M_{X_i}(t) = M_X(t), \forall_i = 1, 2, 3, 4$$

وان

$$M_X(t) = Ee^{tX} = \int_0^\infty e^{tx}.e^{-x}dx = (1-t)^{-1}, t < 1$$

فاذن

$$M_Y(t) = Ee^{tY} = E\,e^{t(X_1+X_2+X_3+X_4)}$$

$$= E \prod_{i=1}^{4} e^{tX_i} = \prod_{i=1}^{4} Ee^{tX_i} = \prod_{i=1}^{4} M_X(t)$$

$$= \prod_{i=1}^{4} (1-t)^{-1} = (1-t)^{-4}.$$

مثال (٣٣) : افرض ان $X_1, X_2, \ldots, X_k$ متغيرات عشوائية مستقلة تسلك وفق دالة توزيع مشترك معين . وان $M(t_1, t_2, \ldots, t_k)$ تمثل الدالة المولدة لعزوم هذا التوزيع . برهن ان

$$\psi(t_1, t_2, \ldots, t_k) = \ln M(t_1, t_2, \ldots, t_k) = \sum_{i=1}^{k} \ln M_{X_i}(t_i)$$

البرهان : حيث ان المتغيرات $X_1, X_2, ..., X_k$ مستقلة فذلك يعني ان

$$M(t_1, t_2, ..., t_k) = \prod_{i=1}^{k} M_{X_i}(t_i)$$

فاذن

$$\psi(t_1, t_2, ..., t_k) = \ln \prod_{i=1}^{k} M_{X_i}(t_i)$$

$$= \sum_{i=1}^{k} \ln M_{X_i}(t_i)$$

٤ ـ ٥ : متراجحة كوشي ـ شوارتز

Cauchy - schwartz inequality

ليكن X_1, X_2 متغيرين عشوائيين بدالة كتلة احتمالية مشتركه $P(x_1, x_2)$ او دالة كثافة احتمالية مشتركة $f(x_1, x_2)$ وافرض ان عزوم التوزيع المشترك موجودة ، عندئذ $(EX_1X_2)^2 \leq (EX_1^2).(EX_2^2)$

البرهان : لتكن $g(Z) = E(X_1 - ZX_2)^2$ دالة ذات قيمة حقيقية بدلالة Z وهي دالة غير سالبة طالما ان $(X_1 - ZX_2)^2 \geq 0$ لجميع قيم X_1, X_2, Z وهذا يعني ان $g(Z) = E(X_1 - ZX_2)^2 \geq 0$ ، الان

$$g(Z) = EX_1^2 + Z^2EX_2^2 - 2ZEX_1X_2$$

واضح ان $g(Z)$ تمثل شكلاً تربيعياً بدلالة Z . وبايجاد المشتقة الاولى لدالة $g(Z)$ نسبة الى Z نحصل على :

$$g'(Z) = 2ZEX_2^2 - 2EX_1X_2$$

وبجعل $g'(Z) = 0$ نحصل على

$$ZEX_2^2 - EX_1X_2 = 0 \rightarrow Z = \frac{EX_1X_2}{EX_2^2}$$

وحيث ان $g''(Z) = 2EX_2^2 > 0$ فذلك يعني ان $g(Z)$ تمتلك نهاية صغرى عندما $Z = \frac{EX_1X_2}{EX_2^2}$ عليه فان :

$$\text{Min } g(Z) = g\left(Z = \frac{EX_1X_2}{EX_2^2} \right)$$

$$= EX_1^2 + \left(\frac{EX_1X_2}{EX_2^2} \right)^2 . EX_2^2 - 2\left(\frac{EX_1X_2}{EX_2^2} \right) . EX_1X_2 \geq 0$$

اي ان

$$(EX_1^2)(EX_2^2) - (EX_1X_2)^2 \geq 0$$

فاذن

$$(EX_1X_2)^2 \leq (EX_1^2)(EX_2^2)$$

واذا كان X_1 مستقل تصادفياً عن X_2 فذلك يعني ان

$$(EX_1)^2 . (EX_2)^2 \leq (EX_1^2)(EX_2^2)$$

مثال (٣٤) : لمعطيات المثال (٧) تحقق من متراجحة كوشي ـ شوارتز.

الحل : واضح من معطيات هذا المثال ان

$$EX_1^2 = \frac{114}{21}, \quad EX_2^2 = \frac{19}{7}, \quad EX_1X_2 = \frac{72}{21}$$

عليه فان .

$$(EX_1X_2)^2 = \left(\frac{72}{21} \right)^2 = 11 \cdot 755101$$

وان

$$(EX_1^2) . (EX_2^2) = \left(\frac{114}{21} \right)\left(\frac{19}{7} \right) = 14 \cdot 734693$$

فاذن نستنتج ان

$$(EX_1X_2)^2 < (EX_1^2)(EX_2^2)$$

مثال (٣٥) : لمعطيات المثال (٨) تحقق من متراجحة كوشي ـ شوارتز .
الحل : واضح من معطيات هذا المثال ان :

$$EX_1^2 = EX_2^2 = \frac{5}{12}, EX_1X_2 = \frac{1}{3}$$

عليه فان :

$$(EX_1X_2)^2 = \frac{1}{9} = 0{\cdot}111111$$

وان

$$(EX_1^2)(EX_2^2) = \frac{5}{12} \cdot \frac{5}{12} = \frac{25}{144} = 0{\cdot}1736111$$

فاذن نستنتج ان

$$(EX_1X_2)^2 < (EX_1^2)(EX_2^2)$$

(تمارين الفصل الرابع)

٤ ـ ١ : لتكن $P(x_1, x_2) = \frac{x_1 + 2x_2}{c}$; $x_1 = 0, 1, 2, x_2 = 1, 2, 3$ تمثل الدالة الاحتمالية المشتركة للمتغيرين X_1, X_2 . يطلب اجراء مايلي :

أ ـ جد قيمة الثابت c .

ب ـ مخطط الدالة $P(x_1, x_2)$

جـ ـ جد $P_r(X_1 = 1, X_2 \leq 2)$ ، $P_r(X_1 \geq 1, X_2 = 3)$.

د ـ جد الدالة الحدية لكل من X_2, X_1

هـ ـ جد معامل الارتباط البسيط بين X_2, X_1

و ـ التوزيع الشرْطي الى X_1 علماً ان $X_2 = x_2$

ز ـ الوسط الشرْطي والتباين الشرْطي الى X_1 علما ان $X_2 = 2$.

ح ـ تحقق من متراجحة كوشي ـ شوارتز .

٤ ـ ٢ :افرض ان دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 موصوفة بالآتي :

(x_1, x_2):	$(1, -1)$	$(1, 0)$	$(1, 1)$	$(2, -1)$	$(2, 0)$	$(2, 1)$
$P(x_1, x_2)$:	P	3P	4P	3P	2P	P

يطلب اجراء مايلي :

أ ـ جد قيمة P ثم ارسم مخطط الدالة $P(x_1, x_2)$.

ب ـ جد $P_r(X_1 = 1, X_2 \leq 0)$ ، $P_r(X_2 \geq 0)$

جـ ـ جد معامل الارتباط البسيط بين X_1, X_2

د ـ التوزيع الشرْطي الى X_2 علماً ان $X_1 = x_1$

هـ ـ الوسط الشرْطي والتباين الشرْطي الى X_2 علماً ان $X_1 = 2$.

و ـ تحقق من متراجحة كوشي ـ شوارتز

٤ ـ ٣ :افرض ان $P(x_1, x_2, x_3) = \frac{x_1 + 2x_2 + 3x_3}{c}$ ، $x_1 = 0, 1, 2$ ، $x_2 = 1, 2, 3$ ، $x_3 = 1, 2$ تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرات X_1, X_2, X_3 . جد مايلي :

أ ــ قيمة الثابت c

ب ــ $P_r(X_1 \le 2, X_2 = 3, X_3 = 2), P_r(X_1 = 1, X_2 \le 2, X_3 = 2)$.

جـ ــ $P(x_i), i = 1, 2, 3 \quad , P(x_i, x_j); i < j$.

د ــ مصفوفة التباين والتباين المشترك لهذه المتغيرات وكذلك مصفوفة معاملات الارتباط البسيط .

هـ ــ $P(x_1 | x_2, x_3) \quad , P(x_2 | x_1, x_3) \quad , P(x_3 | x_1, x_2)$.

و ــ $P(x_1, x_2 | x_3) \quad , P(x_1, x_3 | x_2) \quad , P(x_2, x_3 | x_1)$.

ز ــ $E(X_1 | x_2, x_3) \quad , E(X_2 | x_1, x_3) \quad , E(X_3 | x_1, x_2)$

$E(X_1^2 | x_2, x_3) \quad , E(X_2^2 | x_1, x_3) \quad , E(X_3^2 | x_1, x_2)$

ح ــ $E(X_1X_2 | x_3) \quad , E(X_1X_3 | x_2) \quad , E(X_2X_3 | x_1)$.

ط ــ $\sigma_{12 \cdot 3} \; , \sigma_{13 \cdot 2} \; , \sigma_{23 \cdot 1}$.

$\rho_{12 \cdot 3} \; , \rho_{13 \cdot 2} \; , \rho_{23 \cdot 1}$.

٤ ــ ٤ : لتكن $f(x_1, x_2) = x_1 + x_2 ; 0 < x_1, x_2 < 1$ ، جد الوسط والتباين الشرطي الى X_1 علماً ان $X_2 = x_2$ جد قيمة الوسط والتباين $x_2 = \frac{1}{2}$.

٤ ــ ٥ : اذا علمت ان $f(x_1 | x_2) = c_1 \frac{x_1^2}{x_2^2}, f(x_2) = c_2 x_2^4 ; 0 < x_1 < x_2 < 1$

جد ما يلي :

أ ــ قيمة C_1, C_2 .

ب ــ دالة التوزيع المشترك الى X_1, X_2 .

جـ ــ دالة التوزيع الشرطي الى X_2 علماً ان $X_1 = x_1$.

د ــ $P_r(0.25 < X_1 < 0.5 | X_2 = 0.6), P_r(X_1 < 0.6, X_2 < 0.8)$.

هـ ــ الدالة التوزيعية المشتركة الى X_1, X_2 ثم احسب $F(0.2, 0.3)$.

و ــ الوسط والتباين الشرطي الى X_1 علماً ان $X_2 = 0.7$.

ز ــ معامل الارتباط البسيط بين X_1, X_2 .

ح ــ الدالة المولدة لعزوم التوزيع الشرطي الى X_1 علما ان $X_2 = x_2$.

٤ ــ ٦ : افرض ان $F(x), f(x)$ تمثلان على التوالي دالة الكثافة الاحتمالية والدالة التوزيعية للمتغير العشوائي X . وافرض ان التوزيع الشرطي الى X علما ان $X > x_0$ ، حيث x_0 ثابت حقيقي هو

$f(x | X > x_0) = f(x) / (1 - F(x_0)) ; x > x_0$ ، تحقق من كون ان

هذا التوزيع الشرْطي يتمتع بخصائص دوال الكثافة الاحتمالية . واذا كانت $f(x) = e^{-x}, x \geq 0$ جد مايلي :

أ ـ التوزيع الشرْطي الى X علماً ان $X > 2$ مع رسم مخطط هذه الدالة .

ب ـ $P_r(X < 5 | X > 2)$, $P_r(X > 3 | X > 2)$

جـ ـ الوسط والتباين الشرْطي الى X علماً ان $X > 2$

د ـ الدالة التوزيعية الى X علماً ان $X > 2$ ثم جد قيمة هذه الدالة عندما $X = 4$ مع رسم مخطط هذه الدالة .

هـ ـ الدالة المولدة لعزوم التوزيع الشرْطي الذي حصلت عليه في (أ) ثم جد $E(X^3 | X > 2)$.

٤ ـ ٧ : افرض ان $f(x)$ تمثل دالة الكثافة الاحتمالية الى X وان σ_x^2 , μ_x يمثلان الوسط والتباين لهذا المتغير . ليكن $Y = a + bX$ حيث ان $a , b \neq 0$ ثابتان حقيقيان . جد

أ ـ قيمة a , b بحيث ان $\sigma_y^2 = 1 , \mu_y = 0$.

ب ـ قيمة ρ_{xy} .

جـ ـ $F(y)$ بدلالة a , b .

٤ ـ ٨ : لتكن $f(x,y) = c(6 - x - y); 0 < x < 2, 2 < y < 4$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X , Y . يطلب اجراء مايلي :

أ ـ قيمة $\sigma_{Y|X}^2 , E(Y | X = x), C$

ب ـ تحقق من ان $EY = E[E(Y|X = x)]$

جـ ـ معامل الارتباط البسيط بين X و Y

د ـ الدالة التوزيعية المشتركة ثم جد قيمة $F(1, 3)$

هـ ـ $P_r(X > 1, Y < 3)$, $P_r(Y > 3 | X = 1)$.

و ـ تحقق من متراجحة كوشي ـ شوارتز .

٤ ـ ٩ : افرض ان X_1 , X_2 متغيران عشوائيان بدالة كثافة احتمالية مشتركة $f(x_1, x_2)$ وان $h(V) = E[(X_1 - \mu_{x_1}) - V(X_2 - \mu_{x_2})]^2$ دالة بدلالة V .

أ ـ جد قيمة V التي تجعل $h(V)$ اقل مايمكن .

ب ـ استناداً الى قيمة V المستخرجة في (أ) استخدم متراجحة كوشي ـ شوارتز لبيان ان $-1 \leq \rho_{x_1 x_2} \leq 1$.

٤ ـ ١٠ : افرض ان $\psi(t_1, t_2) = \ln M(t_1, t_2)$ حيث $M(t_1, t_2)$
تمثل الدالة المولدة لعزوم التوزيع المشترك الى X_1, X_2 . برهن ان

$$\left.\frac{\partial\psi(t_1, t_2)}{\partial t_i}\right]_{t_1=t_2=0} = EX_i, \quad \left.\frac{\partial^2\psi(t_1, t_2)}{\partial t_i^2}\right]_{t_1=t_2=0} = \sigma_i^2,$$

$$\left.\frac{\partial^2\psi(t_1, t_2)}{\partial t_1 \cdot \partial t_2}\right]_{t_1=t_2=0} = \sigma_{12}$$

٤ ـ ١١ : لتكن $f(x_1, x_2) = 12x_1x_2(1 - x_2); 0 < x_1, x_2 < 1$. هل
يمكن القول ان X_1 مستقل تصادفياً عن X_2 ؟ ان كان الجواب بالنفي
احسب قيمة ρ_{12} .

٤ ـ ١٢ : لتكن $M_{Y|X}(t)$ تمثل الدالة المولدة لعزوم التوزيع الشرْطي الى Y علماً
ان $X = x$ برهن : اذا كان X مستقل تصادفياً عن Y فان
$M_{Y|X}(t) = M_Y(t)$.

٤ ـ ١٣ : افرض ان X, Y متغيران عشوائيان بدالة كثافة احتمالية مشتركة .
$f(x, y) = \frac{\pi^2}{8} \sin \frac{\pi}{2}(x + y), 0 < x, y < 1$.

أ ـ هل يمكن القول ان X مستقل تصادفياً عن Y ام انهما غير مرتبطين ؟
ب ـ جد التوزيع الحدي لكل من X ، Y .
جـ ـ جد التوزيع الشرْطي الى X علماً ان $Y = y$ ، ثم جد الوسط والتباين
الى X علماً ان $y = 0.8$.

٤ ـ ١٤ : افرض ان X, Y متغيران عشوائيان يتوزعان وفق دالة توزيع مشترك
وان a, b ثابتان حقيقيان . اشتق صيغة الى $Var(aX + bY)$ بدلالة
ρ_{xy} ثم بين ان

$Var(aX + bY) = (a\sigma_x + b\sigma_y)^2$ عندما $|\rho_{xy}| = 1$.

٤ ـ ١٥ : ليكن X متغير عشوائي بدالة كثافة احتمالية $f(x) = \frac{1}{2\pi}; 0 < x < 2\pi$
وافرض ان $u = \sin x$, $V = \cos x$ هل يمكن القول ان u, V مستقلان
تصادفياً ؟

٤ ـ ١٦ : افرض ان X , Y متغيران عشوائيان مستقلان وان μ_x , σ_x^2 يمثلان الوسط والتباين الى X وان μ_y , σ_y^2 الوسط والتباين الى Y . هل يمكن القول ان $Z = X + Y$ مستقل تصادفياً عن $V = X - Y$؟ ان كان الجواب بالنفي اشتق صيغة لمعامل الارتباط بين V , Z

٤ ـ ١٧ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x)=1 \; ; 0 < x < 1$. ليكن $V = \cos 2\pi x \; , Z = \sin 2\pi x$. برهن ان $Var(Z + V) = Var(Z) + Var(V)$ هل ان الاستنتاج السابق يدعونا للقول ان Z مستقل تصادفياً عن V .

٤ ـ ١٨ : لتكن $f(x_1 , x_2) \lambda^2 e^{-\lambda x_2} ; 0 < x_1 < x_2 < \infty$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1 , X_2 . جد الدالة المولدة لعزوم هذا التوزيع المشترك ومن خلال هذه الدالة جد :

أ ـ الوسط والتباين لكل متغير .

ب ـ معامل الارتباط البسيط بينهما .

جـ ـ $M_{x_2|x_1}(t_2) \, , M_{x_1|x_2}(t_1)$.

د ـ على ضوء استنتاجك في (جـ) جد الوسط والتباين الشرطي الى X_1 علماً ان $X_2 = x_2$ وكذلك الى X_2 علماً ان $X_1 = x_1$.

٤ ـ ١٩ : افرض ان Y , X متغيران عشوائيان مستقلان . تحقق من صحة كل مما يلي

أ ـ $$E\left(\frac{X}{Y}\right) = \frac{EX}{EY} \; , \; Var\left(\frac{X}{Y}\right) = \frac{Var(X)}{Var(Y)}$$

ب ـ $$Var(X - Y) = Var(X) - V(Y)$$

جـ ـ $$\sqrt{Var(X + Y)} = \sqrt{Var(X)} + \sqrt{Var(Y)}$$

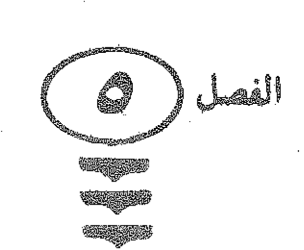
٤ ـ ٢٠ : افرض ان $X_1, X_2 \dots, X_k$ متغيرات عشوائية مستقلة بحيث ان

$Var(X_i) = \sigma^2, EX_i = \mu \quad \forall i = 1, 2, \dots, k$ وافرض ان

$$\sum_{i=1}^{k} a_i = 1 \quad Y = \sum_{i=1}^{k} a_i X_i$$

جد الوسط والتباين الى Y





التوزيعات المتقطعة النظرية

الفصل الخامس

التوزيعات المتقطعة النظرية

لقد تركز اهتمامنا في الفصول السابقة على دراسة التوزيعات الاحتمالية بشكل عام كنماذج رياضية احتمالية ، من حيث خصائصها والامور ذات العلاقة بالتوزيع (كدوال توليد العزوم ، المتوسط ، التباين ، ... الخ) في هذا الفصل سوف نركز اهتمامنا على دراسة بعض عوائل التوزيعات المتقطعة المعلمية(*) فقط ذات الاهمية التطبيقية في النظرية الاحصائية وذلك من خلال استعراض دالة الكتلة الاحتمالية للتوزيع مع عرض لاهم خصائصه .

٥ ـ ١ : التوزيع المنتظم المتقطع

Discrete uniform distribution

يعد هذا التوزيع ابسط التوزيعات النظرية المتقطعة ، ويستخدم هذا التوزيع وبشكل غير مباشر في تلك التجارب التي تتصف نتائجها بكونها ذات نفس الفرصة في الوقوع (كعملية سحب عينة عشوائية بسيطة من مجتمع) . ويعرف هذا التوزيع على النحو الآتي : يقال ان المتغير العشوائي X يتوزع وفق دالة التوزيع المنتظم المتقطع اذا كانت دالة الكتلة الاحتمالية لهذا المتغير تأخذ الشكل التالي :

$$P(x) = P(x;N) = \frac{1}{N} \;;\; x = 1, 2, \ldots, N$$

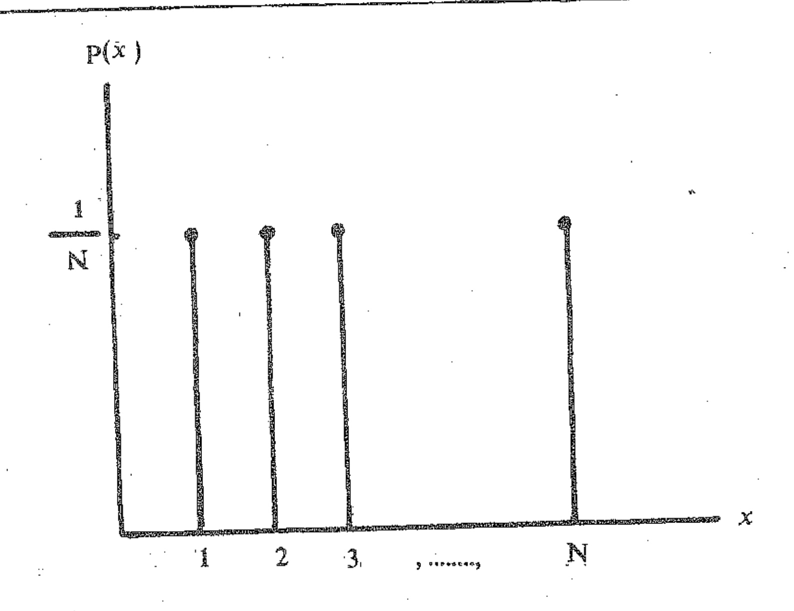
$$= 0 \qquad \text{other wise}$$

حيث N عدد موجب صحيح تمثل معلمة التوزيع . فاذا كانت $N = 6$ فذلك يعني تحديد احد افراد عائلة التوزيع المنتظم المتقطع المشخص بقيمة $N = 6$ وعندئذ

$$P(x;6) = \frac{1}{6} \;;\; x = 1, 2, \ldots, 6$$

(*) ،يقصد بالتوزيع المعلمي بانه ذلك التوزيع الاحتمالي الذي تتضمن دالته ثوابت معينة تسمى معالم التوزيع والتي من شأنها تحديد احد افراد تلك العائلة من خلال تخصيص قيمة عددية parameters لتلك المعلمة او المعالم . علماً ان هنالك توزيعات لاتتضمن ثوابت من هذا النوع تسمى توزيعات لامعلمية .

ان تسمية هذا التوزيع بتوزيع منتظم متقطع ناجمة عن كون قيمة الكتلة الاحتمالية المقترنة باي عنصر من عناصر فضاء العينة للمتغير X ثابتة ومساوية الى $\frac{1}{N}$ وكأن عناصر فضاء المتغير X هي حوادث ذات نفس الفرصة في الوقوع Equally likely events وكما هو موضح في الشكل (٥ ـ ١) :



الشكل (٥ ـ ١) ، مخطط لدالة توزيع منتظم متقطع .

وللسهولة فاننا سوف نرمز للمتغير العشوائيّ الذي يتوزع وفق دالة هذا التوزيع بالشكل X ~ Du (N) وتقرأ كالآتي : المتغير العشوائي X يتوزع "~" كتوزيع منتظم متقطع "Du" بالمعلمة N .

ان مجموع الكتل الاحتمالية المقترنة بعناصر فضاء المتغير X يجب ان يكون مساويا للواحد الامر الذي سمح لنا اطلاق تسمية دالة كتلة احتمالية على هذا التوزيع وذلك واضح من خلال الآتي :

$$\sum_{x=1}^{N} P(x) = \sum_{x=1}^{N} \frac{1}{N} = \frac{1}{N} . N = 1$$

٥ ـ ١ ـ ١ : الدالة التوزيعية Distribution function

سبق وان عرفنا الدالة التوزيعية لمتغير عشوائي X بانها قيمة الاحتمال المتراكم لغاية قيمة معطاة للمتغير X مثل x . وهذا يعني ان :

$$F(x) = P_r(X \leq x) = \sum_{u=1}^{x} P(u) = \sum_{u=1}^{x} \frac{1}{N} = \frac{x}{N} \;; x: 1, 2, \ldots N$$

ويتضح من هذه الدالة ان :

$$F(0) = 0 \quad , \quad F(N) = 1$$

٥ ـ ١ ـ ٢ :الوسط والتباين لقيم X. Mean and variance of X.

ان الوسط لقيم X في هذا التوزيع هو

$$\mu_x = EX = \sum_{x=1}^{N} x \cdot \frac{1}{N} = \frac{1}{N} \sum_{x=1}^{N} x$$

لكن $\sum_{x=1}^{N} x$ يمثل مجموع سلسلة اعداد طبيعية وقيمة هذا المجموع هي $\frac{N(N+1)}{2}$

فاذن

$$\mu_x = \frac{1}{N} \cdot \frac{N(N+1)}{2} = \frac{N+1}{2}$$

كذلك فان تباين قيم X في هذا التوزيع هو :

$$V(X) = EX^2 - (EX)^2 = \frac{1}{N} \sum_{x=1}^{N} x^2 - \left(\frac{N+1}{2} \right)^2$$

لكن $\sum_{x=1}^{N} x^2$ يمثل مجموع مربعات اعداد طبيعية وقيمة هذا المجموع هي :

$\frac{N(N+1)(2N+1)}{6}$ فاذن

$$V(X) = \frac{1}{N} \cdot \frac{N(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2 = \frac{N^2-1}{12}$$

$$= \frac{N-1}{6} \cdot \mu_x$$

ويتضح من الصيغة الاخيرة ان :

$$V(X) \leq \mu_x \text{ for } x \leq 7$$
$$> \mu_x \text{ for } x > 7$$

٥ ـ ١ ـ ٢ : الدالة المولدة للعزوم حول نقطة الاصل .

حسب تعريف الدالة المولدة لعزوم X حول نقطة الاصل (اذا كانت موجودة) فان هذا التوزيع يمتلك دالة مولدة لعزومة وهي :

$$M_X(t) = Ee^{tX} = \frac{1}{N} \sum_{x=1}^{N} e^{tx} = \frac{1}{N} \sum_{x=1}^{N} Z^x , Z = e^t$$

$$= \frac{1}{N}(Z + Z^2 + \dots + Z^N) = \frac{Z}{N}(1 + Z + \dots + Z^{N-1})$$

لكن المجموع داخل القوس الاخير يمثل حدود متوالية هندسية نهائية اساسها مساوٍ الى Z وان مجموع حدودها هو :

$$\sum_{x=0}^{N-1} Z^x = \frac{1-Z^N}{1-Z}$$

عليه فان

$$M_X(t) = \frac{Z}{N} \cdot \frac{1-Z^N}{1-Z} = \frac{e^t(1-e^{Nt})}{N(1-e^t)}$$

فاذن

$$M_X(t) = \frac{e^t(e^{Nt} - 1)}{N(e^t - 1)} \; ; t > 0$$

وبنفس الاسلوب الموضح اعلاه يمكن البيان ان الدالة المميزة لهذا التوزيع هي :

$$\phi(t) = \frac{e^{it}(e^{Nit} - 1)}{N(e^{it} - 1)}$$

٥ ــ ١ ــ ٤ : امثلة

مثال (١) : افرض ان $X \sim Du(8)$ عندئذٍ

$$P(x) = \frac{1}{8} \; ; x = 1, 2, \ldots, 8$$

$$F(x) = \frac{x}{8} \; ; x = 1, 2, \ldots, 8$$

$$EX = \frac{8 + 1}{2} = 4{\cdot}5 \;\; , V(X) = \frac{8^2 - 1}{12} = \frac{63}{12}$$

$$P_r(X \leq 4) = F(4) = 0{\cdot}5 \; . \; P_r(X \geq 3) = 1 - F(2) = 0{\cdot}75$$

مثال (٢) : اذا علمت ان $X \sim Du(N)$ جد الوسط والتباين الى $Y = a + bX$ حيث ان a, b ثابتان حقيقيان

الحل :

$$EY = a + bEX$$

لكن

$$EX = \frac{N + 1}{2}$$

فاذن

$$EY = a + \frac{b(N + 1)}{2}$$

كذلك فان

$$V(Y) = b^2 V(X)$$

لكن

$$V(X) = \frac{N^2 - 1}{12}$$

فاذن

$$V(Y) = \frac{b^2(N^2 - 1)}{12}$$

تمارين عن التوزيع المنتظم المتقطع

٥ ـ ١ : افرض ان $X \sim Du(7)$ جد مايلي :

أ ـ الوسط والتباين لقيم X

ب ـ $P_r(X > \mu_x)$ ، $P_r(\mu_x - \sigma_x \leq X \leq \mu_x + \sigma_x)$

جـ ـ الوسط والتباين الى $Y = 2 + 3X$

٥ ـ ٢ : افرض ان $X \sim Du(N)$ جد اقرب عدد صحيح الى N بحيث ان

$P_r(X \leq \mu_x) = 0.56$

٥ ـ ٣ : ليكن $X \sim Du(N)$ وتطلب الامر اجراء بتر في هذا التوزيع من خلال حذف القيم $M < N, x = 1, 2, \ldots, M$ جد دالة الكتلة الاحتمالية للمتغير X بعد اجراء عملية البتر ثم حدد اسم التوزيع الذي ستحصل عليه . ماهي قيمة الوسط والتباين لهذا التوزيع .

٥ ـ ٤ : اذا كان $X \sim Du(N)$ وان $M_X(t)$ تمثل الدالة المولدة لعزوم X . جد الدالة المولدة التراكمية $K_X(t)$ ثم بين ان $K'_X(0) = \mu_x$ وان $K''_X(0) = \sigma^2_x$. جد $M_X(\ln t)$

٥ ـ ٢ : توزيع برنولي Bernoulli distribution
(محاولات برنولي) (Bernoulli trials)

تأمل تجربة فحص بطارية جافة واحدة (محاولة واحدة) لبيان مدى مطابقتها للمواصفات المحددة من قبل الجهة المنتجة . واضح ان نتيجة الفحص هي اما « البطارية مطابقة للمواصفات » او « البطارية ليست مطابقة للمواصفات » . وعلى فرض ان X يشير الى مطابقة البطارية للمواصفات المحددة فذلك يعني « نجاح المحاولة » . ليكن P يمثل احتمال نجاح المحاولة فان $q = 1 - P$ سوف يمثل احتمال فشلها .

عليه يمكن تخصيص قيمتين ممكنتين فقط للمتغير X هما (0) عند فشل المحاولة و (1) عند نجاحها . ووفق هذا التصور يقال ان المتغير العشوائي X يتوزع وفق دالة توزيع برنولي اذا كانت الكتلة الاحتمالية المقترنة بالعنصر x هي :

$$P(x) = P(x,p) = P^{x}(1 - P)^{1-x} \quad ; x = 0, 1$$
$$= 0 \qquad \text{otherwise}$$

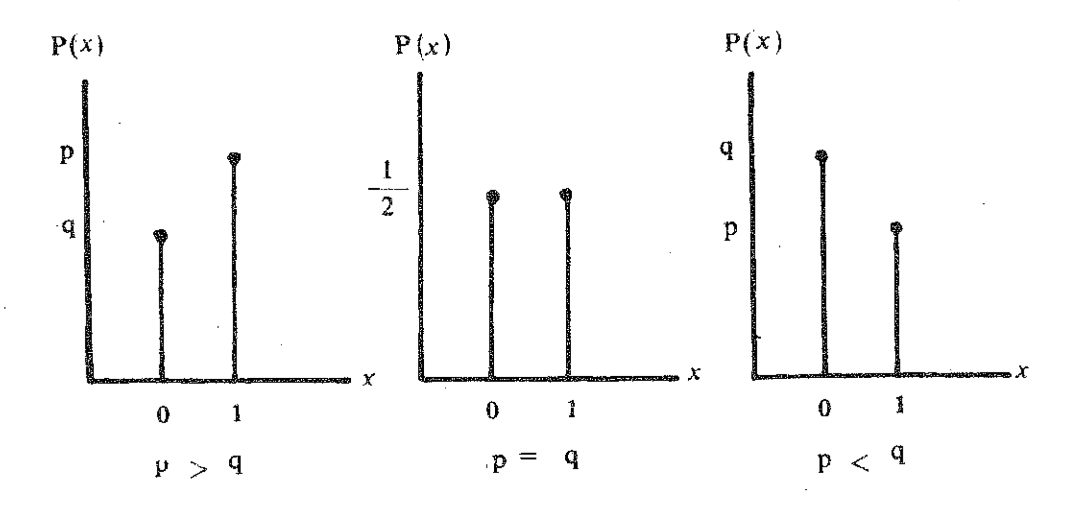
حيث P تمثل معلمة هذا التوزيع بحيث ان $0 < P < 1$. وبالرموز فان $X \sim Ber(P)$. والشكل (٥ ـ ٢) يوضح دالة الكتلة الاحتمالية لهذا التوزيع .

كذلك يتضح مما تقدم ان حاصل جمع الكتلة الاحتمالية المقترنة بالعنصر (0) مع الكتلة الاحتمالية المقترنة بالعنصر (1) مساو للواحد . اي ان :

$$P(x = 0) + P(x = 1) = (1 - P) + P = 1$$

كذلك يلاحظ عندما $P = q$ فان التوزيع في هذه الحالة يمثل توزيعاً منتظماً متقطعاً معرفاً بالدالة $P(x) = \frac{1}{2} \; ; x = 0, 1$. ان الدالة التوزيعية لتوزيع برنولي وببساطة هي :

$$F(x) = q \quad ; x \leq 0$$
$$= 1 \quad ; x \leq 1$$



الشكل (٥ ـ ٢) ، توضيح لدالة توزيع برنولي .

٥ ـ ٢ ـ ١ : الوسط والتباين لتوزيع برنولي .

ان الوسط للمتغير X في توزيع برنولي هو :

$$\mu_x = EX = \sum_x xP(x) = \sum_x xP^x(1-P)^{1-x}$$

$$= (0)(1-P) + (1)(P) = P$$

وهذا يعني ان الوسط لهذا التوزيع هو احتمال نجاح المحاولة . اما تباين X فهو

$$V(X) = EX^2 - (EX)^2$$

لكن

$$EX^2 = \sum_x x^2 . P^x(1-P)^{1-x} = (0)^2(1-P) + (1)^2 . P = p$$

فاذن

$$V(X) = P - P^2 = Pq$$

اي ان تباين X يمثل احتمال نجاح المحاولة مضروب باحتمال فشلها .
ويلاحظ ان $V(X) = \mu_x . q$، وحيث ان $0 < q < 1$ فان $V(X) < \mu_x$.

٥ ـ ٢ ـ ٢ : الدالة المولدة لعزوم توزيع برنولي .

ان دالة توليد عزوم توزيع برنولي حول نقطة الاصل هي :

$$M_X(t) = \sum_x e^{tx} \cdot P^x (1-p)^{1-x}$$

$$= (1-P) + Pe^t = q + Pe^t$$

ويلاحظ من هذه الدالة ان

$$M_X'(0) = M_X''(0) = \ldots = M_X^{(r)}(0) = P$$

اي ان

$$EX^r = P \quad \forall r = 1, 2, \ldots$$

كذلك ووفق نفس الاسلوب اعلاه يمكن ملاحظة ان الدالة المميزة هي :

$$\phi(t) = q + Pe^{it}$$

٥ ـ ٢ ـ ٣ : أمثلة

مثال (١) : افرض ان $X \sim Ber(0{\cdot}6)$. عندئذٍ

$$P(x; 0{\cdot}6) = (0{\cdot}6)^x \cdot (0{\cdot}4)^{1-x} ; x = 0, 1$$

$$EX = P = 0{\cdot}6 , V(X) = Pq = (0{\cdot}6)(0{\cdot}4) = 0{\cdot}24$$

مثال (٢) : اذا علمت ان $X \sim Ber(P)$ ماهو التوزيع الاحتمالي للمتغير $Y = 1 - X$

الحل :

$$P(x) = P^x (1-P)^{1-x} ; x = 0, 1$$

الان اذا كان $x = 0$ فان $y = 1$ واذا كان $x = 1$ فان $y = 0$ وان $X = 1 - Y$ عليه فان

$$P(y = 1 - x) = P^{1-y} \cdot (1-P)^y$$

$$= q^y \cdot (1-q)^{1-y} , y = 0, 1$$

وهذا يعني ان $Y = 1 - X \sim Ber(q)$.

مثال (٣) : اذا علمت ان $X \sim Ber(P)$. جد الوسط والتباين الى $Y = a + bX$.
حيث ان a, b ثابتان حقيقيان .

الحل :

$$EY = a + bEX = a + bP \quad , EX = P$$
$$V(Y) = b^2V(X) = b^2Pq \quad , V(X) = Pq$$

تمارين عن توزيع برنولي

٥ ـ ٥ : افرض ان $X \sim Ber(0.95)$. جد مايلي :
$P(x)$, μ_x , $V(X)$, $M_X(t)$

٥ ـ ٦ : جد الوسط والتباين لتوزيع برنولي باستخدام صيغة الدالة المولدة للعزوم والدالة المميزة .

٥ ـ ٧ : افرض ان $P(x_1, x_2) = P^{x_1 + x_2} . (1 - P)^{2 - x_1 - x_2}$, $x_1, x_2 = 0, 1$
تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 . بين ان التوزيع الحدي للمتغير X_1 وكذلك للمتغير X_2 هو توزيع برنولي .

جد التوزيع الشرطي للمتغير X_1 علماً ان $X_2 = x_2$.

٥ ـ ٨ اذا علمت ان $X \sim Ber(0.7)$. جد الوسط والتباين للمتغير $Y = 2 - 4X$.

٥ ـ ٣ : توزيع ثنائي الحدين Binomial distribution

يعتبر توزيع ثنائي الحدين احد التوزيعات المتقطعة ذات اهمية تطبيقية كبيرة في الحياة العملية وخصوصاً في موضوع الرقابة على جودة الانتاج وموضوع اختبارات النسب والنسب المئوية . ويعد العالم James Bernoulli (١٦٥٤ ـ ١٧٠٥) مكتشف هذا التوزيع عام ١٧٠٠ وتم نشر انجازه هذا عام ١٧١٣ بعد وفاته بثمان سنوات . ويمكن اعتبار هذا التوزيع حالة اكثر عمومية لتوزيع برنولي عندما يكون عدد المحاولات اكثر من محاولة واحدة .

على سبيل المثال عند فحص $n \geq 1$ بطارية جافة كعينة مختارة من احدى وجبات الانتاج فان نتيجة الفحص المختبري قد تبين عدم وجود بطارية معيبة او هناك بطارية واحدة معيبة او اثنتان وهكذا . وهذا يعني ان هنالك n محاولة مستقلة (المحاولة هنا تمثل عملية فحص بطارية واحدة كل مرة) . ووفق هذا التصور يمكن اشتقاق دالة توزيع ثنائي الحدين وكما يلي : بفرض ان p يمثل احتمال نجاح المحاولة (احتمال الحصول على بطارية غير معيبة) وان هذا الاحتمال ثابت من محاولة لاخرى . فان احتمال فشل المحاولة (احتمال الحصول على بطارية معيبة) هو $q = 1 - p$ وان احتمال الفشل سيكون هو الآخر ثابت من محاولة لاخرى بسبب فرض ثبات p . واذا رمزنا للبطارية غير المعيبة (الجيدة) بالرمز g وللبطارية المعيبة بالرمز d . واننا بصدد حساب احتمال ان x من البطاريات غير معيبة (اي ان $n - x$ معيبة) وان نتائج الفحص المختبري لهذه العينة كانت على سبيل المثال :

$$g, g, g, d, g, d, d, d, g, g, \ldots, d, g, d$$

عندئذٍ فان احتمال الحصول على x بطارية غير معيبة و $(n - x)$ بطارية معيبة من بين n بطارية (محاولة مستقلة) هو

$$P_r(g, g, g, d, g, d, d, d, g, g, \ldots, d, g, d) =$$
$$P_r(g) . P_r(g) . P_r(g) . P_r(d) . P_r(g) . P_r(d) \ldots P_r(d) . P_r(g) . P_r(d)$$

وحيث ان في اية محاولة كان الفرض $P_r(g) = P$, $P_r(d) = 1 - P = q$ فاذن

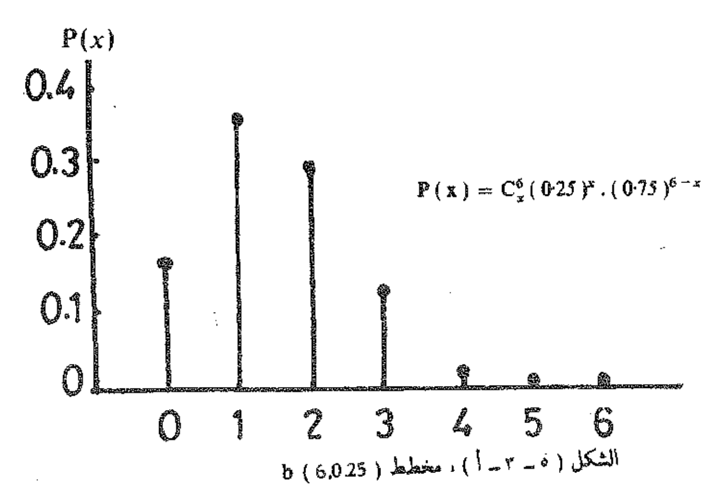
$$P_r(g, g, g, d, g, d, \ldots, d, g, d) = P . P . P . q . P . q \ldots . q . P . q$$
$$= \underbrace{P . P . P \ldots . P}_{\text{حدود عددها } x} . \underbrace{q . q . q \ldots . q}_{\text{حدود عددها } n-x} = P^x . q^{n-x}$$

وحيث انه يمكن الحصول على x بطارية غير معيبة من بين n بطارية ، $x \le n$ من خلال C_x^n طريقة متاحة وان احتمال وقوع اية طريقة منها هو $P^x . q^{n-x}$ فذلك يعني ان احتمال الحصول على x بطارية غير معيبة من بين n بطارية اياً كان ترتيب نتائج الفحص هو $P_r(x) = C_x^n P^x . q^{n-x}$ ان هذه الكتلة الاحتمالية المقترنة بالعنصر x تسمى دالة توزيع ثنائي الحدين . وقد جاءت تسمية التوزيع بتوزيع « ثنائي الحدين » بسبب اننا في كل محاولة نتخذ قراراً ذا حدين ومن النوع « نجاح المحاولة » او « فشل المحاولة » ، « جيد » او « غير جيد » ، « مطابق » او « غير مطابق » وغيرها من الالفاظ المماثلة . واستناداً لما سبق يمكن تعريف توزيع ثنائي الحدين على النحو التالي : يقال ان المتغير العشوائي X يتوزع وفق دالة توزيع ثنائي الحدين اذا كان دالة الكتلة الاحتمالية لهذا المتغير تأخذ الشكل التالي :

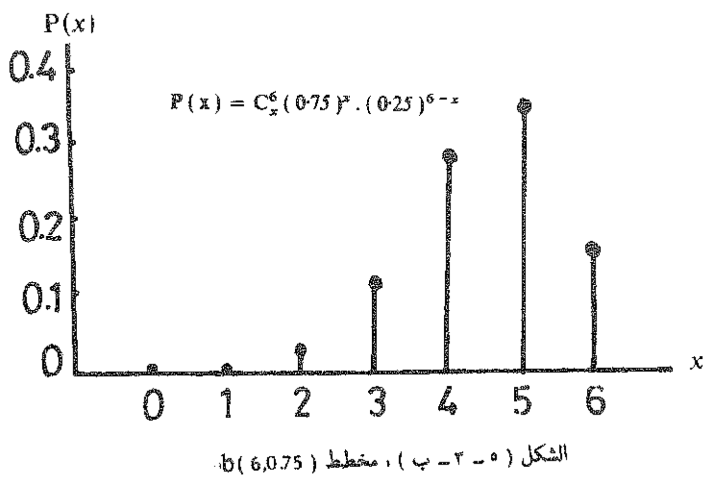
$$
\begin{aligned}
P(x) = P(x ; n, P) &= C_x^n . P^x q^{n-x} ; x = 0, 1, 2, \dots, n \\
&= 0 \qquad , \text{other wise}
\end{aligned}
$$

حيث n , p تمثلان معلمتي التوزيع بحيث ان n عدد موجب صحيح (عدد المحاولات) وان P تمثل احتمال نجاح المحاولة حيث $q = 1 - P , 0 < P < 1$

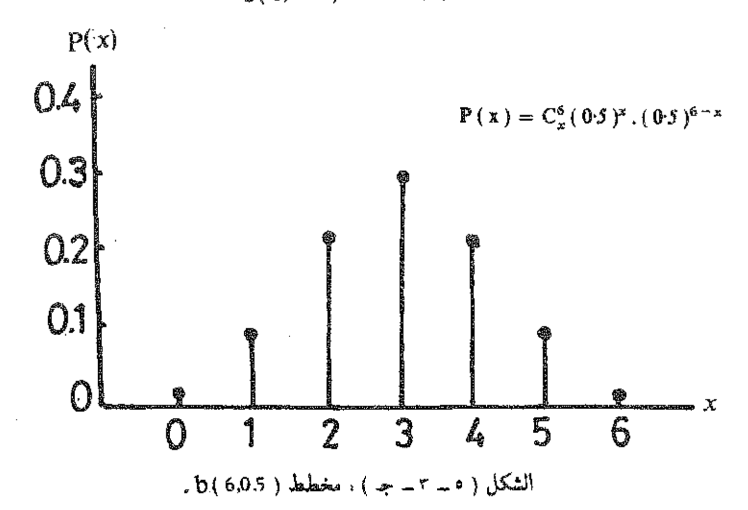
وللسهولة في الترميز يقال ان $X \sim b(n, P)$. ان تخصيص قيمة لكل من n , P تعني تحديد احد افراد عائلة توزيعات ثنائي الحدين . والشكل (٥ ــ ٣) يوضح مخطط دالة توزيع ثنائي الحدين .



الشكل (٥ ـ ٣ ـ أ) ، مخطط (b (6,0.25



الشكل (٥ ـ ٣ ـ ب) ، مخطط (b(6,0.75



الشكل (٥ ـ ٣ ـ جـ) ، مخطط (b(6,0.5 .

ويلاحظ من الشكل (٥ ـ ٣) مايلي :

١ ـ اذا كانت $q > P$ فذلك يعني ان التوزيع ذو التواء موجب (لاحظ الشكل أ)

٢ ـ اذا كانت $q < P$ ذلك يعني ان التوزيع ذو التواء سالب (لاحظ الشكل ب)

٣ ـ اذا كانت $q = P$ فذلك يعني ان التوزيع متماثل (لاحظ كل جـ)

مع ملاحظة انه كلما كانت P قريبة من الصفر (عند ثبات قيمة n) فذلك يعني ان q تكون قريبة من الواحد وهذا يعني زيادة شدة الالتواء الموجب. في حين كلما كانت P قريبة من الواحد (عند ثبات قيمة n) فذلك يعني ان q تكون قريبة من الصفر وهذا يعني زيادة شدة الالتواء السالب. وكلما كانت P قريبة من q وكلاهما يقترب من 0.5 فذلك يعني الاقتراب من حالة التماثل. ويمكن بيان ان مجموع الكتل الاحتمالية المقترنة بعناصر فضاء العينة للمتغير X مساوٍ للواحد وكالاتي :

$$\sum_{x=0}^{n} P(x\,;n\,,P) = \sum_{x=0}^{n} C_x^n P^x \cdot q^{n-x}$$

ولاي عددين مثل a , b فان نظرية ثنائي الحدين Binomial theorem تنص بما يلي :

$$(a+b)^n = \sum_{k=0}^{n} C_k^n a^k \cdot b^{n-k}$$

وبفرض ان $a = p\,, b = q$ عندئذٍ

$$(P+q)^n = \sum_{k=0}^{n} C_k^n P^k \cdot q^{n-k}$$

لكن $P + q = 1$ فاذن $(P+q)^n = 1$ طالما ان n عدد معرف. وهذا يعني ان

$$\sum_{x=0}^{n} C_x^n P^x \cdot q^{n-x} = 1$$

ان هذا الاثبات الى جانب كون دالة هذا التوزيع غير سالبة ووحيدة القيمة يدعونا للقول بانها دالة كتلة احتمالية.

٥ ـ ٣ ـ ١ : الدالة التوزيعية

ان دالة التراكم الاحتمالي لتوزيع ثنائي الحدين هي :

$$F(x) = P_r(X \leq x) = \sum_{k=0}^{x} C_k^n P^k q^{n-k}$$

وهذه الصيغة تبين قيمة الاحتمال المتراكم حتى قيمة معينة من قيم X هي x . ويلاحظ وجود صعوبة في التعامل مع هذه الدالة تطبيقياً وعلى الاخص عندما تكون قيمة n كبيرة فمثلاً لو كان $X \sim b(50, 0.8)$ وتطلب الامر حساب $F(25)$ فذلك يعني انه يتوجب حساب قيمة $P(x)$ عند قيم $x = 0, 1, 2, \ldots, 25$ ومن ثم اجراء عملية جمع هذه النتائج أي ان ذلك يتطلب التعويض في الدالة $P(x)$ ستة وعشرون مرة وبعد ذلك يتم حساب قيمة $F(25)$ كحاصل جمع لنتائج التعويض . وعلى أية حال يمكن القضاء على هذه الصعوبة من خلال صياغة برنامج على حاسب الكتروني بلغة بيسك أو فورتران من شأنه حساب التراكم الاحتمالي لأية قيمة للمعلمة n . وهنالك جداول خاصة بهذا التوزيع تبين الاحتمال المتراكم لغاية قيمة من قيم المتغير X وعند قيم مختلفة لكل من p,n . هذه الجداول استند في حسابها على ما يسمى بـ « تكامل بيتا الناقص **Incomplete beta integral** والذي سيتم عرض شكل آخر له يسمى « تكامل بيتا الكامل أو تكامل بيتا وذلك في الفقرة (٦ ـ ٥) ومن خلال هذا الشكل سيتم تعريف تكامل بيتا الناقص المعطى هنا بالصيغة التالية :

$$I_x = (n-x) C_x^n \int_0^q t^{n-x-1} . (1-t)^x \, dt$$

والمطلوب بيانه أن

$$F(x) = \sum_{k=0}^{x} C_k^n p^k q^{n-k} = I_x$$

الان نعود لتكامل بيتا الناقص وباجراء التكامل بطريقة التجزئة من خلال الفرض ان

$$u = (1-t)^x , dv = t^{n-x-1} dt$$

فاذن

$$du = x(1-t)^{x-1}.(-dt), V = \frac{t^{n-x}}{n-x}$$

وهذا يعني ان

$$\int_0^q t^{n-x-1}(1-t)^x dt = \frac{p^x . q^{n-x}}{n-x} + \frac{x}{n-x}\int_0^q t^{n-x}.(1-t)^{x-1} dt$$

فاذن

$$I_x = C_x^n p^x . q^{n-x} + xC_x^n \int_0^q t^{n-x}.(1-t)^{x-1} dt$$

$$= P(x;n,p) + I_{x-1}$$

ولو عدنا مرة اخرى لاجراء التكامل بطريقة التجزئة لحل I_{x-1} بنفس الاسلوب الموضح في المرة الاولى لحصلنا على

$$I_{x-1} = p(x-1;n,p) + I_{x-2}$$

فاذن

$$I_x = p(x;n,p) + p(x-1;n,p) + I_{x-2}$$

واذا استمر الحال اجراء التكامل بطريقة التجزئة فاننا سوف نحصل على

$$I_x = p(x;n,p) + p(x-1;n,p) + \dots + p(1,n,p) + p(0;n,p)$$

$$= \sum_{k=0}^{x} p(x;n,p) = \sum_{k=0}^{x} C_x^n p^x . q^{n-x} = F(x)$$

وهذا يعني انه يمكن حساب الاحتمال المتراكم حتى قيمة معينة من قيم X مثل x باستخدام تكامل بيتا الناقص . والجدول (1) الموضح في الملحق (ب) يبين قيم الكتل الاحتمالية المقترنة بعناصر فضاء X عند قيم مختارة لكل من n , p أي ان هذا الجدول يعرض قيم $p(x,n,p)$ والتي من خلالها يمكن اجراء عملية الجمع لغاية قيمة معطاة الى X مثل x بهدف الحصول على $F(x)$.

ويلاحظ ان هذا الجدول مصمم لقيم $p \leq 0.5$ واذا تطلب الامر حساب كتلة احتمالية لتوزيع ثنائي الحدين معرف بـ $p > 0.5$ فانه يمكن استخدام الخاصية التالية التي توضح ان :

$$P(x; n, p) = p(n - x, n, q)$$

فمثلاً لو كان المطلوب هو حساب $p(8; 10, 0.6)$ وهذا الاحتمال غير معرف في الجدول (١) ملحق (ب) ، الا انه ممكن الحساب وفق مايلي :

$$p(8; 10, 0.6) = p(2; 10, 0.4) = 0.1209$$

٥ ـ ٣ ـ ٢ : الوسط والتباين لتوزيع ثنائي الحدين :

ان الوسط لقيم X في توزيع ثنائي الحدين هو $\mu_x = np$ وان تباين هذا المتغير هو $\sigma_x^2 = npq$

البرهان :

$$\mu_x = \sum_{x=0}^{n} x \cdot C_x^n p^x \cdot q^{n-x}$$

$$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!.(n-x)!} \cdot p^{x-1} \cdot q^{n-x}$$

الان بفرض ان $n' = n - 1, y = x - 1$ عندئذٍ

E(x)

$$\mu_x = np \sum_{y=0}^{n'} C_y^{n'} \cdot p^y \cdot q^{n-y} = np, Y \sim b(n', p)$$

كذلك فان $\sigma_x^2 = EX^2 - (Ex)^2$ لكن

$$EX^2 = \sum_{x=0}^{n} x^2 \cdot C_x^n p^x \cdot q^{n-x}$$

وبجعل $x^2 = x(x-1) + x$ نحصل على :

$$EX^2 = \sum_{x=2}^{n} x(x-1) C_x^n p^x . q^{n-x} + np$$

$$= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!.(n-x)!} p^{x-2} . q^{n-x} + np$$

وبفرض ان $n^* = n-2, y = x-2$ عندئذٍ

$$EX^2 = n(n-1)p^2 \sum_{y=0}^{n^*} C_y^{n^*} p^y . q^{n^*-y} + np$$

$$= n(n-1)p^2 + np, y \sim b(n^*, p)$$

عليه فان

$$\sigma_x^2 = n(n-1)p^2 + np - n^2p^2 = npq$$

ويتضح من هذه الصيغة ان $\sigma_x^2 = \mu_x . q$ وهذا يعني ان تباين X هو اقل من وسطه .

٥ ـ ٣ ـ ٣ : الدالة المولدة للعزوم حول نقطة الاصل .

ان توزيع ثنائي الحدين يمتلك دالة مولدة للعزوم حول نقطة الاصل معطاة بالصيغة : $M_X(t) = (q + pe^t)^n$

البرهان :

$$M_X(t) = Ee^{tX} = \sum_{x=0}^{n} e^{tx} . C_x^n p^x . q^{n-x}$$

$$= \sum_{x=0}^{n} C_x^n (pe^t)^x . q^{n-x}$$

وحسب نظرية ثنائي الحدين ولاي عددين مثل a,b فان :

$$(a+b)^n = \sum_{k=0}^{n} C_x^n a^k . b^{n-k}$$

وبجعل $b = q , a = pe^t$ فان :

$$M_X(t) = (pe^t + q)^n$$

واضح ان :

$$M_X(0) = (p+q)^n = 1 , K_X(t) = ln\, M_X(t) = n\, ln(pe^t + q).$$

ووفق نفس الاسلوب يمكن البيان ان $\phi(t) = (pe^{it} + q)^n$

٥ ـ ٢ ـ ٤ : صيغة التراجع Recurrence formula

ان صيغة التراجع بشكل عام تعني حساب الكتلة الاحتمالية المقترنة بالعنصر $x+1$ اذا علمت الكتلة الاحتمالية المقترنة بالعنصر السابق x . وفيما يلي اشتقاق هذه الصيغة في توزيع ثنائي الحدين ، ان

$$p(x) = C_x^n p^x q^{n-x} , p(x+1) = C_{x+1}^n p^{x+1} . q^{n-(x+1)}$$

فاذن :

$$\frac{p(x+1)}{p(x)} = \frac{C_{x+1}^n . p^{x+1} . q^{n-(x+1)}}{C_x^n . p^x . q^{n-1}} = \frac{(n-x)}{(x+1)} . \frac{p}{q}$$

وهذا يعني ان

$$p(x+1) = \left[\frac{(n-x)}{(x+1)} . \frac{p}{q} \right] p(x)$$

ان لهذه الصيغة اهمية تطبيقية كبيرة في حساب الكتل الاحتمالية المقترنة بعناصر فضاء X دون اللجوء للتعويض في الدالة p (x). فبمجرد ايجاد قيمة p (0) فانه يتم تلقائياً حساب الكتل الاحتمالية اللاحقة باستخدام هذه الصيغة .

فمثلاً ، اذا كان $X \sim b(n, p)$ فان $p(0) = q^n$ وان

$$p(1) = \frac{np}{q} \cdot p(0) = npq^{n-1}, \; p(2) = \frac{n(n-1)}{2} p^2 \cdot q^{n-2}, \ldots$$

٥ ـ ٣ ـ ٥ : خاصية الجمع Additive property

اذا كانت $X_1, X_2, \ldots, X_k$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim b(n_i, p)$ فان

$$Y = \sum_{i=1}^{k} X_i \sim b\left(\sum_{i=1}^{k} n_i, p \right)$$

البرهان :

بفرض ان Y له دالة مولدة للعزوم، فان

$$M_Y(t) = Ee^{tY} = Ee^{t \sum_{i=1}^{k} X_i} = E \prod_{i=1}^{k} e^{tX_i}$$

وبما ان X_i متغيرات مستقلة فان

$$M_Y(t) = \prod_{i=1}^{k} Ee^{tX_i}$$

لكن Ee^{tX_i} يمثل تعريف الدالة المولدة لعزوم X_i حول نقطة الاصل وان $X_i \sim b(n_i, p)$ لذلك فان

$$M_{X_i}(t) = (pe^t + q)^{ni}, \; i = 1, 2, \ldots, k$$

عليه فان :

$$M_Y(t) = \prod_{i=1}^{k} (pe^t + q)^{ni} = (pe^t + q)^{\sum_{i=1}^{k} n_i}$$

والدالة الاخيرة تمثل تعريف المولدة لعزوم المتغير العشوائي Y الذي يتضح بانه يتوزع وفق دالة توزيع ثنائي الحدين بالمعلمتين $\left(\sum_{i=1}^{k} n_i , P\right)$ بسبب ان الدالة المولدة لعزوم التوزيع (اذا كانت موجودة) فهي دالة وحيدة وتشخص التوزيع الذي اشتقت منه . فاذن نستنتج ان

$$\sum_{i=1}^{k} X_i \sim b\left(\sum_{i=1}^{k} n_i , P\right)$$

ملاحظة :

اذا كانت $X_i \sim b(n_i , P_i), i=1,2,\ldots,k$ متغيرات مستقلة فان توزيع $\sum_{i=1}^{k} X_i$ سوف لن يكون توزيع ثنائي الحدين . ويطلب من القاريء برهنة ذلك ووفق نفس الاسلوب المشار اليه اعلاه .

٥ ـ ٣ ـ ٦ : أمثلة

مثال (١) : افرض ان $X \sim b\left(8, \frac{1}{4}\right)$. عندئذ

$$1- \quad p(x) = C_x^8 \left(\frac{1}{4}\right)^x \cdot \left(\frac{3}{4}\right)^{8-x} ; x = 0, 1, 2, \ldots, 8 .$$

$$2- \quad F(x) = \sum_{k=0}^{x} C_x^8 \left(\frac{1}{4}\right)^k \cdot \left(\frac{3}{4}\right)^{8-k}$$

$$= (8-x)\, C_x^8 \int_0^{3/4} t^{7-x} . (1-t)^x \, dt .$$

$$3- \quad \mu_x = np = 8\left(\frac{1}{4}\right) = 2 , \quad \sigma_x^2 = npq = 8\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)$$

$$= 1.5$$

$$4 - \quad M_X(t) = \left(\frac{1}{4} e^t + \frac{3}{4} \right)^8$$

$$5 - \quad P_r(X > 1) = 1 - P_r(X \leq 1) = 1 - [P(0) + P(1)]$$

$$= 1 - \left[\left(\frac{3}{4} \right)^8 + 8 \left(\frac{1}{4} \right) \left(\frac{3}{4} \right)^7 \right]$$

$$= 0{\cdot}633$$

مثال (٢) : افرض ان

$$X_3 \sim b\left(10, \frac{1}{3}\right), X_2 \sim b\left(8, \frac{1}{3}\right), X_1 \sim b\left(6, \frac{1}{3}\right)$$

وان هذه المتغيرات مستقلة تصادفياً عندئذ :

$$1 - \quad Y = X_1 + X_2 + X_3 \sim b\left(24, \frac{1}{3}\right)$$

$$2 - \quad p(y) = C_y^{24} \left(\frac{1}{3} \right)^y \cdot \left(\frac{2}{3} \right)^{24-y} ; y = 0, 1, 2, \ldots, 24$$

$$3 - \quad \mu_y = 24 \left(\frac{1}{3} \right) = 8 = \mu_{x_1} + \mu_{x_2} + \mu_{x_3}$$

$$4 - \quad \sigma_y^2 = 24 \left(\frac{1}{3} \right) \left(\frac{2}{3} \right) = \frac{16}{3} = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2$$

مثال (٣) : رميت ثمان قطع من النقود المتجانسة 200 مرة . في كل رمية تم تسجيل عدد الصور الظاهرة X وكانت نتائج هذه التجربة كما موضح في التوزيع التكراري التالي :

عدد الصور	:(x)	8	7	6	5	4	3	2	1	0
التكرار	:(f)	6	9	18	38	40	45	22	12	10

يطلب توفيق توزيع ثنائي الحدين لهذه البيانات .

الحل :

ان احتمال ظهور الصورة للقطعة الواحدة هو $p = \frac{1}{2}$

وهذا يعني ان احتمال ظهور الكتابة هو $q = \frac{1}{2}$

فاذن احتمال ظهور X صورة في رمية واحدة هو

$$p(x) = C_x^8 \left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{8-x}$$

$$= C_x^8 \left(\frac{1}{2}\right)^8 ; x = 0, 1, ..., 8$$

وعلى ضوء هذه الدالة نبدأ بحساب الكتل الاحتمالية المقترنة بعناصر فضاء X . وهنا يمكن استخدام صيغة التراجع في حسابها أو الحصول على هذه الكتل من جدول توزيع ثنائي الحدين (جدول ١ ملحق ب) عندما $n = 8$ ، $p = \frac{1}{2}$ هذه الكتل هي :

x :	0	1	2	3	4	5	6	7	8
P(x) :	0·0039	0·0312	0·1094	0·2188	0·2734	0·2188	0·1094	0·0312	0·0039

بعد ذلك يتم حساب ما يسمى بـ التكرار المتوقع(*) Expected frequency أو التكرار النظري من خلال ضرب الكتلة الاحتمالية المقترنة بالعنصر x بعدد مرات الرمي البالغة 200 . أي ان $E.f = 200 . P(x)$.

والجدول التالي يبين التكرارات المشاهدة والتكرارات المتوقعة المقابلة لقيم x في هذه التجربة .

عدد الصور x :	0	1	2	3	4	5	6	7	8
التكرار المشاهد :	10	12	22	45	40	38	18	9	6
التكرار المتوقع :	1	6	22	44	54	44	22	6	1

(*) يقصد بالتكرار المتوقع بانه عدد مرات تكرار قيم x في المجتمع الاحصائي لتلك التجربة بفرض ثبات عدد مرات تكرار التجربة ففي المثال اعلاه فان عدد مرات تكرار تجربة الرمي كان 200 مرة .

مع ملاحظة ان مجموع التكرارات المشاهدة يجب ان يكون مساوياً لمجموع لتكرارات المتوقعة (لماذا) . في هذا المثال ثم استنتاج قيمة P على ضوء معطيات التجربة (معلوم ان احتمال ظهور صورة هو $\frac{1}{2}$) . اما في حالة عدم توفر معلومات عن P عندئذٍ يستوجب الامر تقدير قيمة لها وكما هو موضح بالمثال (٤) .

مثال (٤) : لجدول التوزيع التكراري التالي يطلب توفيق ثنائي الحدين.

قيم (x) :	5	4	3	2	1	0
التكرار (f) :	4	13	35	31	61	6

الحل : واضح ان $n = 5$ شكن P غير معلومة . وحيث اننا بصدد توفيق توزيع ثنائي الحدين لهذه البيانات عليه نجعل الوسط الحسابي للتوزيع التكراري اعلاه مساوياً الى nP . اي $\bar{x} = nP$ ثم نجد $\bar{x}$ وفق الصيغة :

$$\bar{x} = \frac{\sum_{x=0}^{5} xf_x}{\sum_{x=0}^{5} f_x} = \frac{300}{150} = 2$$

فاذن :

$$\bar{x} = np = 2 \rightarrow 2 = 5p \rightarrow p = \frac{2}{5} = 0{\cdot}4$$

عليه فان.

$$P(x) = C_x^5 (0{\cdot}4)^x (0{\cdot}6)^{5-x} ; x = 0, 1, 2, 3, 4, 5 .$$

ومن جدول توزيع ثنائي الحدين نلاحظ ان الكتل الاحتمالية المقترنة بعناصر فضاء X عندما $n = 5$ ، $P = 0{\cdot}4$ هي :

x :	0	1	2	3	4	5
P(x) :	0·0778	0·2592	0·3456	0·2304	0·0768	0·0102

عليه فان التكرارات المتوقعة يتم الحصول عليها من خلال ضرب مجموع التكرارات المشاهدة بالكتل الاحتمالية P(x) وكما هو موضح بالجدول التالي :

قيم X :	0	1	2	3	4	5
التكرار المشاهد :	6	61	31	35	13	4
التكرار المتوقع :	12	39	52	35	11	1

مثال (٥) : رميت ثمان قطع من النقود مرة واحدة . ماهو احتمال الحصول على صورة واحدة على الاقل . وما هو احتمال الحصول على سبع صور تماماً . جد متوسط عدد الصور في هذه التجربة .

الحل : واضح هنا ان هنالك متغير عشوائي يمثل عدد الصور الظاهرة في رمي ثمان قطع من النقود وليكن هذا المتغير هو X.

وكما هو معلوم فان احتمال الحصول على صورة هو $P=\frac{1}{2}$ وذلك يعني ان احتمال الحصول على كتابة هو $q=\frac{1}{2}$ (اي فشل المحاولة) .

نستنتج مما تقدم ان $X \sim b\left(8, \frac{1}{2}\right)$ فاذن :

$$P(x) = C_x^8\left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{8-x} ; x = 0, 1, 2, \ldots, 8$$

عليه فان احتمال الحصول على صورة واحدة على الاقل هو :

$$P_r(X \geq 1) = 1 - P_r(X < 1) = 1 - P_r(X = 0)$$

$$= 1 - C_0^8\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{8-0} = 1 - \left(\frac{1}{2}\right)^8 = 0{\cdot}9961$$

وان احتمال الحصول على سبع صور تماماً هو :

$$P_r(X = 7) = C_7^8\left(\frac{1}{2}\right)^7 \cdot \left(\frac{1}{2}\right)^1 = 8\left(\frac{1}{2}\right)^8 = 0{\cdot}0313$$

كذلك فان متوسط عدد الصور في هذه التجربة هو :

$$EX = nP = 8\left(\frac{1}{2}\right) = 4$$

مثال (٦) : اذا كان $X \sim b(n, P)$

أ ـ برهن ان $Y = n - X \sim b(n, q)$

ب ـ جد $\frac{Cov(X, n - X)}{E(X - EX)^2 / n}$

جـ ـ جد

الحل :

(أ) سوف نستخدم اسلوب الدالة المولدة للعزوم حول نقطة الاصل في ايجاد توزيع $Y = n - X$. وبفرض ان $M_Y(t)$ موجودة فذلك يعني ان :

$$M_Y(t) = Ee^{tY} = Ee^{t(n-X)} = e^{nt}Ee^{-tX} = e^{nt} . M_X(-t)$$

وحيث ان $X \sim b(n, P)$ فذلك يعني ان :

عليه فان

$$M_X(-t) = (Pe^{-t} + q)^n$$
$$M_Y(t) = (e^t)^n . (Pe^{-t} + q)^n = [e^t(Pe^{-t} + q)]^n$$
$$\therefore M_Y(t) = (P + qe^t)^n$$

والدالة الاخيرة تمثل الدالة المولدة لعزوم توزيع ثنائي الحدين بالمعلمتين n, q ، وهذا يعني ان $Y = n - X \sim b(n, q)$

(ب) :

$$\begin{aligned} Cov(X, n - X) &= E(X - EX)(n - X - E(n - X)) \\ &= E(X - nP)(n - X - n + nP) \\ &= -E(X - nP)^2 \\ &= -nPq = -V(X) \end{aligned}$$

(جـ) :

$$E(X - EX)^2 / n = E(X - nP)^2 / n = nPq / n = Pq$$

تمارين عن توزيع تنائي الحدين

٥ ـ ٩ : افرض ان $X_1 \sim b(15, 0.8)$ مستقل عن $X_2 \sim b(10, 0.8)$ ، وافرض ان $Y = X_1 + X_2$ ،يطلب اجراء مايلي :

أ ـ جد التوزيع الاحتمالي للمتغير Y .

ب ـ ارسم مخطط الدالة $P(y)$

جـ ـ جد الوسط والتباين للمتغير Y

هـ ـ جد $P_r(\mu_y - \sigma_y \leq Y \leq \mu_y + \sigma_y), P_r(Y \geq \mu_y)$

: اذا علمت ان 6% من المصابيح المنتجة في مصنع معين معيبة . اختيرت عينة عشوائية قوامها 18 مصباح من احدى وجبات الانتاج . ما هو احتمال :

أ ـ عدم وجود مصباح معيب في هذه العينة .

ب ـ وجود على الاكثر ثلاثة مصابيح معيبة .

جـ ـ وجود على الاقل عشرة مصابيح غير معيبة .

د ـ ما هو متوسط عدد البطاريات المعيبة في هذه العينة .

٥ ـ ١١ : اذا علمت ان $X \sim b(4, P)$ وان $Y \sim b(6, P)$ كذلك فان

$P_r(X \geq 1) = \frac{5}{9}$ جد $P_r(Y > 1)$.

٥ ـ ١٢ : اذا كان $X \sim b(n, P)$ جد $Eq^{n-X}, EP^X, EP^X.q^{n-X}$ ثم بين ان

$EP^X . Eq^{n-X} = (1 - Pq)^{2n}$.

٥ ـ ١٣ : اذا كان $X=K$ يمثل المنوال الوحيد في توزيع $b(n, P)$ برهن ان

$$(n + 1)P - 1 < K < (n + 1)P$$

٥ ـ ٤ : توزيع ثنائي الحدين السالب. Negative binomial dist. والتوزيعات ذات العلاقة به and related distributions.

يسمى هذا التوزيع في بعض الاحيان بـ « توزيع باسكال » نسبة للعالم الرياضي الفرنسي Blaise Pascal (١٦٢٣ ـ ١٦٦٢) . ويعتبر هذا التوزيع واحداً من التوزيعات المتقطعة ذات اهمية تطبيقية في الكثير من المجالات العملية وخصوصاً في العلوم الزراعية وعلم البكتريا . وسمي هذا التوزيع بـ « ثنائي الحدين السالب » بسبب ان دالة كتلته الاحتمالية تمثل الحد العام في مفكوك ثنائي الحدين السالب كما سنلاحظ ذلك لاحقاً .

افرض ان تجربة معينة يمكن تكرارها بعدد كبير من المحاولات المستقلة بحيث ان احتمال نجاح اية محاولة منها مقدار ثابت هو P وافرض ان $P(x; r, p)$ يمثل احتمال الحصول على X محاولة فاشلة التي تسبق حالة الحصول على r محاولة ناجحة من بين $x + r$ محاولة . الان بفرض ان المحاولة الاخيرة كانت محاولة ناجحة ، وهذا يعني ان احتمال هذه المحاولة هو P (احتمال النجاح) وان في المتبقي من المحاولات البالغ عددها $(x + r - 1)$ محاولة سوف يكون هنالك $(r-1)$ محاولة ناجحة احتمال وقوعها سيكون $C^{x+r-1}_{r-1}\, P^{r-1} . q^x$ وعندئذٍ فان احتمال الحصول على x محاولة فاشلة أي $P(x; r, p)$ هو

$$
\begin{aligned}
P(x ;, r, p) &= C^{x+r-1}_{r-1} . P^{r-1} . q^x . P \\
&= C^{x+r-1}_{r-1} . P^r . q^x
\end{aligned}
$$

والشكل الاخير يسمى دالة الكتلة الاحتمالية لتوزيع ثنائي الحدين السالب . وفيما يلي تعريف متكامل لهذا التوزيع ؛ يقال ان المتغير العشوائي x هو ذو توزيع ثنائي الحدين السالب اذا كانت دالة الكتلة الاحتمالية التي يتوزع وفقها هذا المتغير تأخذ الشكل التالي :

$$
P(x ; r, p) = \begin{cases} C^{x+r-1}_{r-1}\, P^r q^x \; ; & x = 0, 1, 2, \ldots \\ 0 & \text{other wise} \end{cases}
$$

حيث r , p معلمتا التوزيع بحيث ان $0 < P < 1, r > 0$ ، $q = 1 - P$. وبالرموز يقال ان $X \sim Nb(r, P)$. كذلك فقد يلاحظ في بعض الاحيان انّ

تكتب دالة هذا التوزيع بالشكل $P(x;r,p) = C_x^{x+r-1}.p^r.q^x$ وذلك بسبب ان $C_{r-1}^{x+r-1} = C_x^{x+r-1}$. كما ويمكن صياغة دالة هذا التوزيع بشكل آخر (دون ان يغير ذلك من مضمون التوزيع) مشتق بالاساس من مفكوك ثنائي الحدين السالب وكما هو مبين بالآتي :

ان

$$C_x^{x+r-1} = \frac{(x+r-1)!}{x!(r-1)!} = \frac{(x+r-1)(x+r-2)\dots(r+1).r(r-1)!}{x!(r-1)!}$$

$$= \frac{r(r+1)(r+2)\dots(x+r-2)(x+r-1)}{x!}$$

واضح ان عدد الحدود المضروبة ببعضها في بسط الصيغة الاخيرة مساوٍ الى X حد فاذن :

$$C_x^{x+r-1} = (-1)^x . \frac{(-r)(-r-1)(-r-2)\dots(-r-x+2)(-r-x+1)}{x!}$$

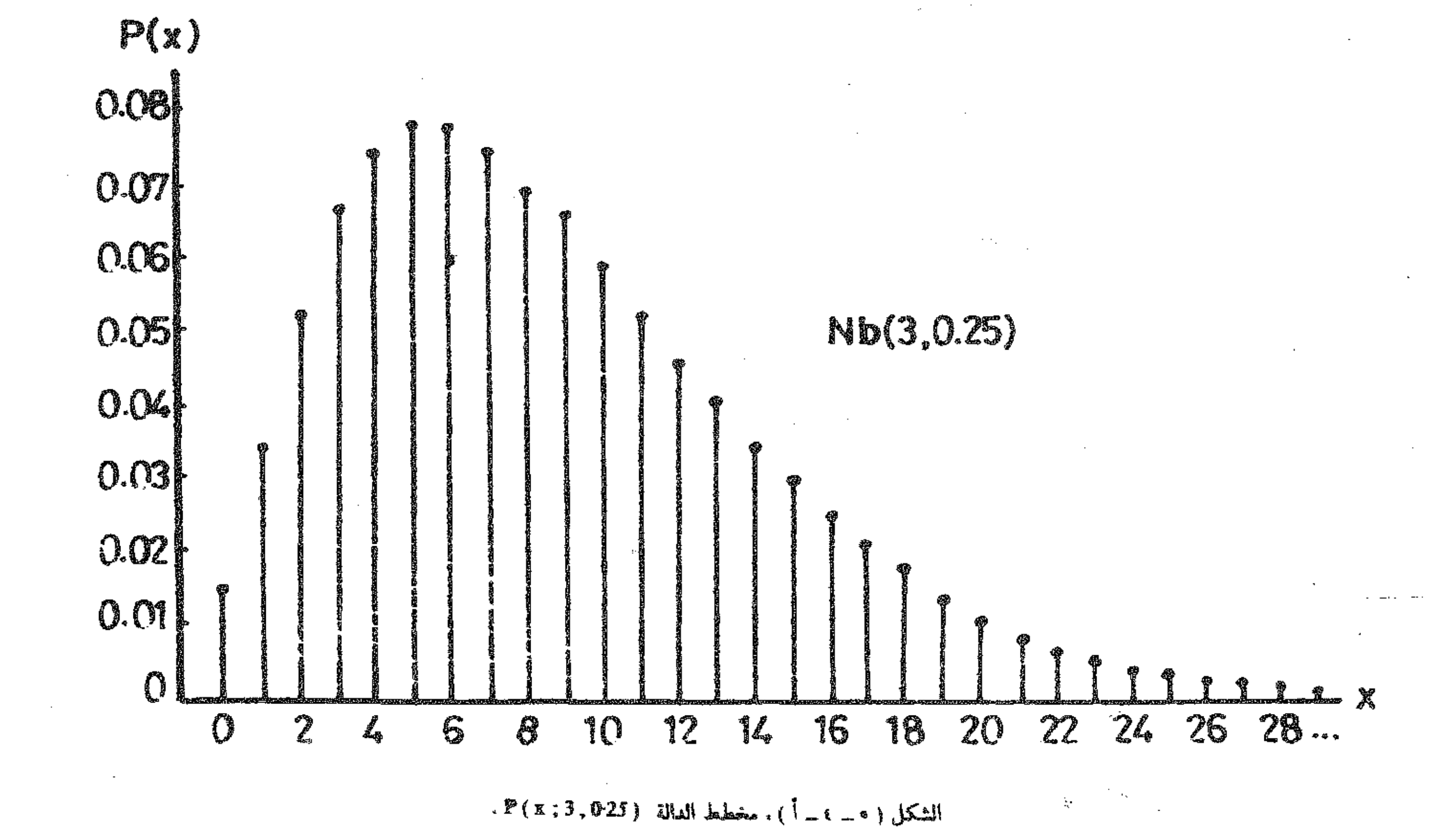
وبضرب البسط والمقام بـ $(-r-x)!$ نحصل على .

$$C_x^{x+r-1} = (-1)^x . \frac{(-r)!}{x!(-r-x)!} = (-1)^x . C_x^{-r}$$

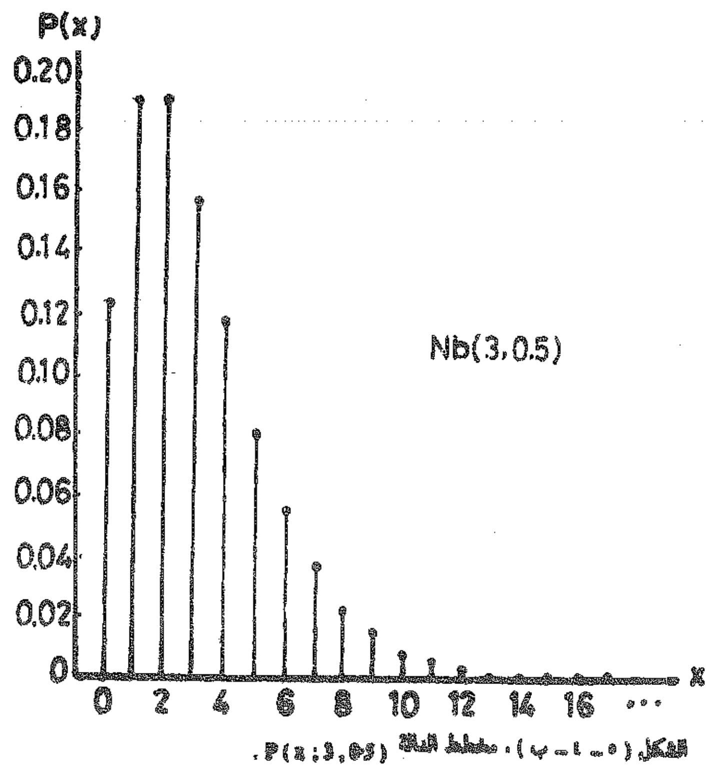
فاذن

$$P(x;r,p) = (-1)^x C_x^{-r} P^r q^x$$
$$= C_x^{-r} P^r (-q)^x \quad ; x = 0, 1, 2, \dots$$

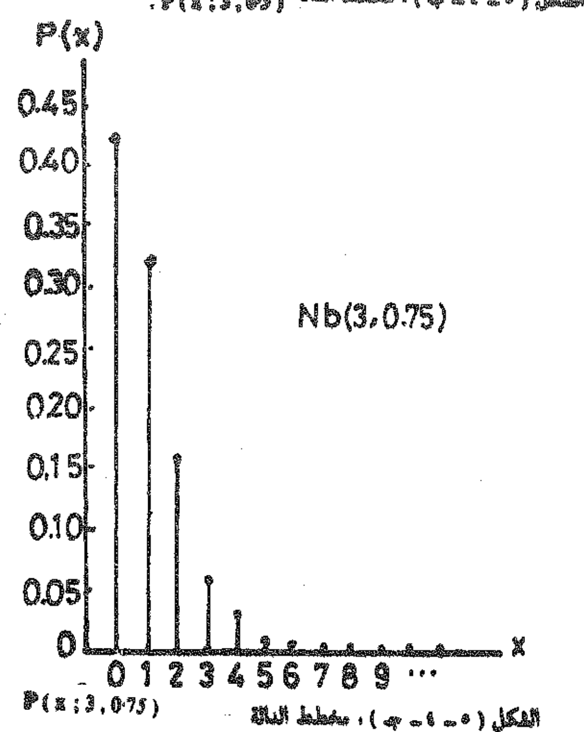
ان الصيغة الاخيرة ما هي الا قيمة الحد ذي التسلسل$(x+1)$في مفكوك ثنائي الحدين السالب اي مفكوك الصيغة $P^r(1-q)^{-r}$ ومن هنا جاءت تسمية هذا التوزيع . والاشكال (٥ ـ ٤) تبين مخطط دالة هذا التوزيع .



الشكل (٥ ـ ٤ ـ أ) ، مخطط الدالة $P(x; 3, 0.25)$.



الشكل (٥ ـ ١ ـ ب)، مخطط الدالة $P(x;3,0.5)$.



الشكل (٥ ـ ١ ـ جـ)، مخطط الدالة $P(x;3,0.75)$

ويمكن بيان ان مجموع الكتل الاحتمالية المقترنة بعناصر فضاء X مساوٍ للواحد دلالة على كون P (x) هي دالة كتلة احتمالية وكالاتي :

$$\sum_{x=0}^{\infty} C^{-r}_{x} P^{r}(-q)^{x} = P^{r} \sum_{x=0}^{\infty} C^{-r}_{x}(-q)^{x}$$

لكن وحسب مفكوك ثنائي الحدين السالب فان

$$\sum_{x=0}^{\infty} C^{-r}_{x}(-q)^{x} = (1-q)^{-r}$$

فاذن

$$\sum_{x=0}^{\infty} C^{-r}_{x} P^{r}(-q)^{x} = P^{r}(1-q)^{-r} = P^{r}.P^{-r} = 1$$

٥ ـ ٤ ـ ١ : الدالة التوزيعية Distribution function

ان الدالة التوزيعية لتوزيع ثنائي الحدين السالب هي :

$$F(x) = P_r(X \leq k) = P^{r} \sum_{x=0}^{k} C^{x+r-1}_{x}.q^{x}$$

علماً انه لا يمكن صياغة هذه الدالة على نحو ابسط . ويلاحظ وجود صعوبة في حساب التراكم الاحتمالي وخصوصاً عندما تكون K كبيرة . لذلك تم التوصل الى عدة صيغ تقريبية اهم هذه الصيغ واكثرها دقة هي تلك المقترحة من قبل **Bartko** عام ١٩٦١ التي تستند في حساب F (x) على جداول التوزيع الطبيعي المعياري (الذي سيأتي ذكره في الفقرة (٦ ـ ٢) وهي $P_r(X \leq k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}u^2} du$. حيث

$$Z = \frac{1}{3}\left[\frac{9k+8}{k+1} - \frac{(9r-1)\left(\frac{rc}{k+1}\right)^{\frac{1}{3}}}{r}\right]\left[\frac{\left(\frac{rc}{k+1}\right)^{\frac{2}{3}}}{r} + \frac{1}{k+1}\right]^{-\frac{1}{2}}$$

وان $C = \frac{q}{P}$ فمثلاً لو تطلب الامر حساب F (3) لتوزيع Nb (10, 0·5) فان قيمة F (3) باستخدام دالة هذا التوزيع هي :

$$F(3) = \left(\frac{1}{2}\right)^{10} \sum_{x=0}^{3} C_x^{x+9} \left(\frac{1}{2}\right)^x = 0{\cdot}0461425$$

في حين ان قيمة F (3) باستخدام تقريب Bartko هي الاتي :

ان c = 1 , k = 3 , r = 10 بسبب ان q = P وبالتعويض عن هذه المعطيات نحصل على Z = − 1·684 . ومن جداول التوزيع الطبيعي المعياري المشار اليه في الفقرة (٦ ـ ٢) نلاحظ ان .

$$F(3) = P_r(X \leq 3) \simeq \int_{-\infty}^{-1{\cdot}684} e^{-\frac{1}{2}z^2} dz = 0{\cdot}0465$$

لاحظ ان مقدار الخطأ المطلق بين الاحتمالين هو 0·0003575 ، وقد اوضح Bartko ان مقدار الخطأ يتضاءل بزيادة قيمة r .

٥ ـ ٤ ـ ٢ : الوسط والتباين Mean and variance

ان الوسط في توزيع ثنائي الحدين السالب هو $\mu_x = rq / P$ وان التباين هو $\sigma_x^2 = rq / P^2$

البرهان :

$$\mu_x = \sum_{x=0}^{\infty} x . C_x^{-r} P^r . (-q)^x$$

$$= rqP^r \sum_{x=1}^{\infty} \frac{(-r-1)!}{(x-1)!(-r-x)!} (-q)^{x-1}$$

$$= rqP^r \sum_{y=0}^{\infty} C_y^{-r^*} (-q)^y ; y = x-1 , r^* = r+1$$

لكن وحسب مفكوك ثنائي الحدين السالب فان :

$$\sum_{y=0}^{\infty} C_y^{-r^*} (-q)^y = (1-q)^{r^*}$$

فاذن

$$\mu_x = rqP^r (1-q)^{-r^*} = rqP^r \cdot P^{-r-1} = \frac{rq}{P}$$

كذلك فان

$$\sigma_x^2 = EX^2 - (EX)^2$$

وان

$$EX^2 = E[X(X-1)+X] = \sum_{x=0}^{\infty} x(x-1) C_x^{-r} P^r (-q)^x + \frac{rq}{P}$$

$$= P^r \sum_{x=2}^{\infty} \frac{(-r)!}{(x-2)!(-r-x)!} (-q)^x + \frac{rq}{P}$$

$$= r(r+1)P^r q^2 \sum_{y=0}^{\infty} C_z^{-r'} (-q)^z + \frac{rq}{P} ; z = x-2, r' = r+2$$

لكن $\sum_{y=0}^{\infty} C_z^{-r'} (-q)^z = (1-q)^{-r'}$ فاذن

$$EX^2 = r(r+1)P^r q^2 (1-q)^{-r-2} + \frac{rq}{P} = \frac{rq(rq+1)}{P^2}$$

فاذن

$$\sigma_x^2 = \frac{rq(rq+1)}{P^2} - \frac{r^2q^2}{P^2} = \frac{rq}{P^2}$$

ويلاحظ من صيغة التباين ان $\sigma_x^2 = \frac{1}{p} \cdot \mu_x$ وحيث ان $\frac{1}{p} > 1$ فذلك يعني ان $\sigma_x^2 > \mu_x$. وهذه تماماً هي عكس ما هو عليه في توزيع ثنائي الحدين حيث لاحظنا حينذاك ان $\sigma_x^2 < \mu_x$

٥ ـ ٤ ـ ٣ : الدالة المولدة للعزوم Moment generating function

ان الدالة المولدة للعزوم حول نقطة الاصل في هذا التوزيع هي $\left(\frac{p}{1-qe^t}\right)^r$

البرهان :

$$M_X(t) = Ee^{tX} = \sum_{x=0}^{\infty} e^{tx} . C_x^{-r} P^r . (-q)^x$$

$$= P^r \sum_{x=0}^{\infty} C_x^{-r} (-qe^t)^x$$

لكن ولاي عدد مثل a فانه وحسب نظرية ثنائي الحدين السالب :

$\sum_{x=0}^{\infty} C_x^{-r} (-a)^x = (1-a)^{-r}$ وبوضع $a = qe^t$ نحصل على

$$M_X(t) = \left(\frac{P}{1-qe^t}\right)^r$$

ووفق نفس الاسلوب المشار اليه اعلاه يمكن ملاحظة ان

$$\phi(t) = \left(\frac{P}{1-qe^{it}}\right)^r$$

كذلك فان

$$K_X(t) = ln\, M_X(t) = r(lnp - ln(1-qe^t))$$

٥ ـ ٤ ـ ٤ : صيغة التراجع Recurrence formula

ان صيغة التراجع في توزيع ثنائي الحدين السالب هي :

$$P(x+1) = \frac{x+r}{x+1} . q . p(x)$$

البرهان :

$$\frac{p(x+1)}{p(x)} = \frac{C^{-r}_{x+1}\, p^r . (-q)^{x+1}}{C^{-r}_{x}\, p^r . (-q)^{x}}$$

وتبسيط هذا المقدار نحصل على :

$$\frac{p(x+1)}{p(x)} = \frac{x+r}{x+1}.q \quad \therefore p(x+1) = \frac{x+r}{x+1}.q.p(x)$$

وكما هو معلوم فان صيغة التراجع مهمة جداً في ايجاد الكتل الاحتمالية المقترنة بعناصر فضاء المتغير X دون اللجوء للتعويض المباشر في الدالة $p(x)$ وانما يستوجب فقط ايجاد الكتلة الاحتمالية المقترنة بالعنصر $x = 0$ ويتم تلقائياً تحديد الكتل الاخرى اللاحقة عن طريق صيغة التراجع .

٥ ـ ٤ ـ ٥ : خاصية الجمع Additive property

لتكن $X_1, X_2, \ldots, X_n$ متغيران عشوائية مستقلة بحيث ان

$X_i \sim Nb(r_i, b), i = 1, 2, \ldots, n$ عندئذٍ $\sum_{i=1}^{n} X_i \sim Nb\left(\sum_{i=1}^{n} r_i, p\right)$

البرهان : افرض ان $M_Y(t)$ تمثل الدالة المولدة لعزوم $Y = \sum_{i=1}^{n} X_i$ فاذن

$$M_Y(t) = Ee^{tY} = Ee^{t(X_1 + X_2 + \cdots + X_N)}$$

$$= E \prod_{i=1}^{n} e^{tX_i}$$

وحيث ان المتغيرات مستقلة فذلك يعني ان

$$M_Y(t) = \prod_{i=1}^{n} Ee^{tX_i}$$

لكن Ee^{tX_i} يمثل تعريف الدالة المولدة لعزوم المتغير $X_i \sim Nb(r_i, p)$ وذلك يعني ان $M_{X_i}(t) = \left(\frac{p}{1 - qe^t} \right)^{r_i}$ عليه فان

$$M_Y(t) = \prod_{i=1}^{n} \left(\frac{p}{1 - qe^t} \right)^{r_i} = \left(\frac{p}{1 - qe^t} \right)^{r}$$

حيث $r = \sum_{i=1}^{n} r_i$ وحيث ان الدالة المولدة للعزوم تشخص التوزيع الاحتمالي الذي اشتقت منه . فاذن نستنتج ان $\sum_{i=1}^{n} X_i \sim Nb\left(\sum_{i=1}^{n} r_i, p \right)$

٥ ـ ٤ ـ ٦ : التوزيع الهندسي Geometric distribution

اذا تم اختيار $r = 1$ في توزيع ثنائي الحدين السالب نحصل على توزيع آخر مهم في التطبيقات الاحصائية وخصوصاً في موضوع الاحصاء السكاني لدى دراسة معدلات نمو السكان ومعدلات الولادات والوفيات هذا التوزيع هو « التوزيع الهندسي » . وهذا يعني ان دالة الكتلة الاحتمالية في التوزيع الهندسي هي :

$$\begin{aligned} p(x, p) &= p \cdot q^x \; ; x = 0, 1, 2, \ldots, 0 < p < 1 \\ &= 0 \qquad \text{other wise} \end{aligned}$$

وبالرموز فان $X \sim G(P)$ وحيث ان هذا التوزيع يمثل حالة خاصة من توزيع ثنائي الحدين السالب لذا فان خصائصه هي نفس خصائص توزيع ثنائي الحدين السالب بمجرد التعويض عن $r = 1$ ماعدا الدالة التوزيعية التي يمكن صياغتها بالاتي

$$F(x) = P_r(X \leq x) = P \sum_{k=0}^{x} q^k$$

لكن المجموع اعلاه يمثل مجموع حدود متوالية هندسية نهائية اساسها مساوٍ الى $q < 1$ وهذا المجموع مساوٍ الى $\frac{1 - q^{x+1}}{1 - q}$ فاذن

$$F(x) = p \cdot \frac{1 - q^{x+1}}{\quad} = 1 - q^{x+1}, x = 0, 1, 2, \ldots$$

اما بقية خصائص هذا التوزيع فهي :

أ ـ ان الوسط والتباين في التوزيع الهندسي هما :

$$\mu_x = \frac{q}{p} , \sigma_x^2 = \frac{q}{p^2}$$

ب ـ ان الدالة المولده لعزوم التوزيع الهندسي هي :

$$M_X(t) = \frac{p}{1 - qe^t}$$

جـ ـ اذا كانت $X_1 , X_2 , \dots , X_l$ متغيرات عشوائية مستقلة بحيث $X_i \sim G(p)$ فان

$\sum_{i=1}^{r} X_i \sim Nb(r,p)$ ونترك برهنة ذلك للقاريء .

٥ ـ ٤ ـ ٧ : توزيع پوليا Polya's distribution

اذا تم اختيار $r = \frac{1}{\beta}$ ، $p = \frac{1}{1+\theta\beta}$ في توزيع ثنائي الحدين السالب

نحصل على توزيع آخر يسمى « توزيع پوليا » دالة كتلته الاحتمالية معطاة بالآتي :

$$p(x;\theta,\beta) = \frac{\prod_{j=0}^{x-1}(1+j\beta)}{x!} . \theta^x \left(\frac{1}{1+\theta\beta} \right)^{x+\frac{1}{\beta}} , x = 0, 1, \dots$$

$$= 0 \qquad \text{other wise}$$

حيث θ,β معلمتا هذا التوزيع بحيث ان $\beta > 0 , \theta > 0$ وبالرموز يقال ان

$X \sim polya(\theta,\beta)$ وان اهم خصائص هذا التوزيع مايلي :

أ ـ ان الوسط والتباين في هذا التوزيع هما :

$$\mu_x = \theta , \sigma_x^2 = \theta(1+\theta\beta)$$

ب ـ ان الدالة المولدة لعزوم التوزيع هي

$$M_X(t) = [1 + \theta\beta(1-e^t)]^{-\frac{1}{\beta}}$$

جـ ـ اذا تم اختيار $\beta = 1$ في توزيع پوليا عندئذٍ نحصل على توزيع هندسي بالمعلمة $p = \frac{1}{1+\theta}$ وذلك واضح من خلال مايلي . بالتعويض عن $\beta = 1$ في دالة توزيع پوليا نلاحظ ان :

$$p(x;\theta,1) = \frac{\prod_{j=0}^{x-1}(1+j)}{x!}\,\theta^{x}\left(\frac{1}{1+\theta}\right)^{x+1}$$

$$= \frac{1\cdot2\cdot3\ldots x}{x!}\cdot\left(\frac{1}{1+\theta}\right)\left(\frac{\theta}{1+\theta}\right)^{x}$$

$$= \left(\frac{1}{1+\theta}\right)\left(\frac{\theta}{1+\theta}\right)^{x},\ x = 0, 1, 2, \ldots$$

وهذه هي دالة توزيع هندسي بالمعلمة $p = \frac{1}{1+\theta}$.

٥ ـ ٤ ـ ٨ : امثلة

مثال (١) : اذا كان $X \sim Nb(6, 0{\cdot}4)$ فان :

$$1-\ \ p(x) = C_{x}^{x+5}(0{\cdot}4)^{6}(0{\cdot}6)^{x};\ x = 0, 1, 2, \ldots$$

$$2-\ \ \mu_x = \frac{rq}{p} = 9\ ,\ \sigma_x^2 = \frac{rq}{p^2} = 22{\cdot}5$$

$$3-\ \ M_X(t) = \left(\frac{0{\cdot}4}{1-0{\cdot}6\,e^{t}}\right)^{6}$$

$$4-\ \ P_r(X \ge 1) = 1 - P_r(0) = 1 - (0{\cdot}4)^{6} = 0{\cdot}995904$$

مثال (٢) : اذا كان $X_1 \sim Nb(7,0.5)$ مستقل عن $X_2 \sim Nb(5,0.5)$ وان $Y = X_1 + X_2$ فان :

$$1 - \quad Y \sim Nb(12, 0.5), \; P(Y) = C_y^{y+11}(0.5)^{y+12}\,; \; y = 0, 1, \ldots$$

$$2 - \quad \mu_y = \frac{12(0.5)}{0.5} = 12, \; \sigma_y^2 = \frac{12(0.5)}{(0.5)^2} = 24$$

مثال (٣) : لجدول التوزيع التكراري التالي يطلب توفيق توزيع ثنائي الحدين السالب .

x :	0	1	2	3	4	5
f_x :	200	140	40	20	8	2

الحل : حتى نتمكن من توفيق توزيع ثنائي الحدين السالب فان ذلك يتطلب تحديد قيمة p, r . وكما يلي :
ان الوسط الحسابي لهذا التوزيع هو :

$$\bar{x} = \frac{\sum_{x=0}^{5} x f_x}{\sum_{x=0}^{5} f_x} = \frac{322}{410} = 0.7853658$$

وان التباين هذا التوزيع هو :

$$S^2 = \frac{\sum_{x=0}^{5} x^2 f_x}{\sum_{x=0}^{5} f_x} - \bar{x}^2 = \frac{658}{410} - (0.7853658)^2 = 0.9880786$$

وبوضع

$$\mu_x = \frac{rq}{p} = 0.7853658, \; \sigma_x^2 = \frac{rq}{p^2} = 0.9880786$$

وحل المعادلتين نسبة الى r , p نحصل على :

$$r = 3.0427249 \quad , p = 0.7948414 \quad , q = 0.2051586$$

ونفرض ان $X \sim Nb\,(3.0427249, 0.7948414)$. وعلى اساس هذا الفرض نبدأ بحساب الكتل الاحتمالية المقترنة بقيم X باستخدام صيغة التراجع .

ان

$$P(0) = (0.7948414)^{3.0427249} = 0.497257$$

فادن

$$P(1) = rq \cdot p(0) = 0.3104082$$

$$P(2) = \frac{r+1}{2} \cdot q \cdot p(1) = 0.1287262$$

$$P(3) = \frac{r+2}{3} \cdot q \cdot p(2) = 0.0443915$$

$$P(4) = \frac{r+3}{4} \cdot q \cdot p(3) = 0.0137582$$

$$P(5) = \frac{r+4}{5} \cdot q \cdot p(4) = 0.0039758$$

عليه فان التكرارات المتوقعة المقابلة لقيم X سوف تنتج من حاصل ضرب الكتل الاحتمالية بمجموع التكرارات اي ان $E.f = P(x)\Sigma f_x$ الموضحة في الجدول الآتي :

قيم (x) :	5	4	3	2	1	0
التكرار المشاهد :	2	8	20	40	140	200
التكرار المتوقع :	2	6	18	53	127	204

تمارين عن توزيع ثنائي الحدين السالب

٥ ـ ١٤ : افرض ان $X \sim Nb(4, 0.3)$ يطلب اجراء ما يلي

أ ـ جد دالة الكتلة الاحتمالية للمتغير X ثم ارسم مخطط هذه الدالة .

ب ـ جد الوسط والتباين للمتغير X

جـ ـ جد العزم الثالث والرابع حول نقطة الاصل باستخدام الدالة المولدة للعزوم .

د ـ جد الوسط والتباين والدالة المولدة لعزوم $Y = 4 + 5X$

هـ ـ جد $P_r(X \geq 1)$, $P_r(X \leq 6)$, $P_r(2 < X \leq 8)$:

و ـ استخدم تقريب Bartko في حساب قيمة تقريبية الى $P_r(X \leq 20)$

٥ ـ ١٥ : اذا كان $X \sim Nb(r, p)$ جد : EP^X , EP^{X-r} , $Eq^r . P^x$ ، $E\left(\frac{X}{r}\right)$, $V\left(\frac{X}{\sqrt{r}}\right)$.

٥ ـ ١٦ : افرض ان $X_1 \sim G(P)$ مستقل عن $X_2 \sim G(P)$: برهن ان التوزيع الشرطي للمتغير X_1 علماً ان $X_1 + X_2 = n$ هو توزيع منتظم متقطع

٥ ـ ١٧ : اشتق صيغة للدالة المولده للعزوم المركزية والداللة المولدة للعزوم العاملية لكل من .

أ ـ توزيع ثنائي الحدين السالب :

ب ـ التوزيع الهندسي .

جـ ـ توزيع پوليا .

٥ ـ ١٨ : برهن ان صيغة العزم العاملي ذا المرتبة k :

أ ـ في التوزيع الهندسي هي $K! \mu_x^k$

ب ـ في توزيع ثنائي الحدين السالب هي $\left(\frac{q}{p}\right)^k . \prod_{j=1}^{k} (r + j - 1)$

جـ ـ في توزيع پوليا هي $\theta^k . \prod_{j=1}^{k} (1 + (j-1)\beta)$

٥ ـ ٥ : التوزيع الهندسي الزائدي

Hypergeometric distribution

افرض ان صندوقاً يحتوي على N كره ، M منها حمراء اللون والبقية N − M سوداء . وافرض انه تم سحب عينة عشوائية قوامها n (بدون ارجاع) من هذا الصندوق عندئذٍ فان احتمال الحصول على X كرة حمراء ضمن هذه العينة هو :

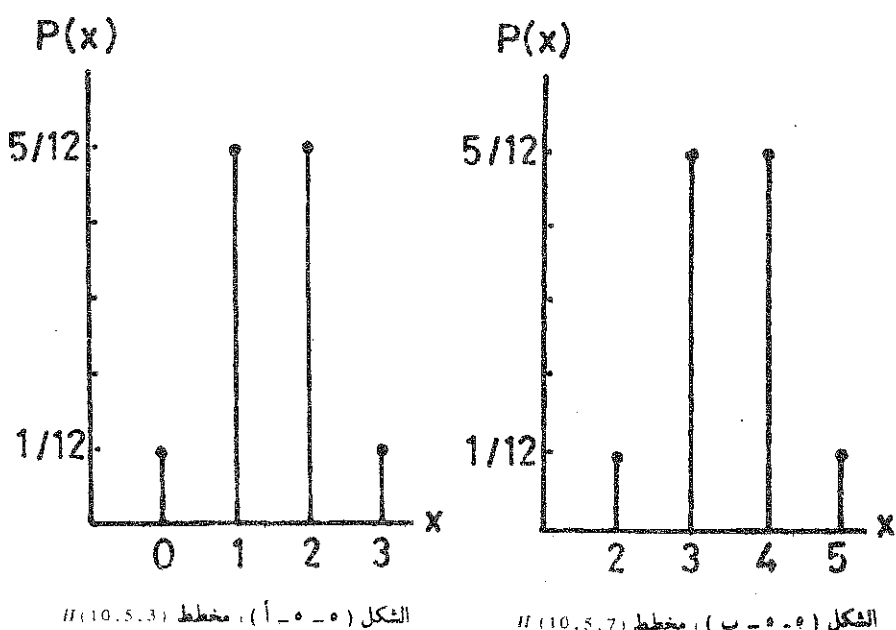
$$P_r(X = x) = \frac{C_x^M \cdot C_{n-x}^{N-M}}{C_n^N}, \; x \leq n$$

ووفق هذا المثال البسيط يمكن تعريف التوزيع الهندسي الزائدي على النحو التالي : يقال ان المتغير العشوائي X هو ذو توزيع هندسي زائدي اذا كانت دالة الكتلة الاحتمالية التي يتوزع وفقها هذا المتغير تأخذ الشكل التالي :

$$P(x; N, M, n) = \frac{C_x^M \cdot C_{n-x}^{N-M}}{C_n^N} \; ; \max(0, n-N+M) \leq x \leq \min(n, M)$$

$$= 0 \quad \text{other wise}$$

حيث N , M , n تمثل معالم هذا التوزيع جميعاً تمثل اعداد موجبة صحيحة بحيث ان M ≤ N وان n ≤ N . وبالرموز نقول ان X ~ H (N , M , n) الاشكال (٥ ـ ٥) توضح مخطط دالة هذا التوزيع :

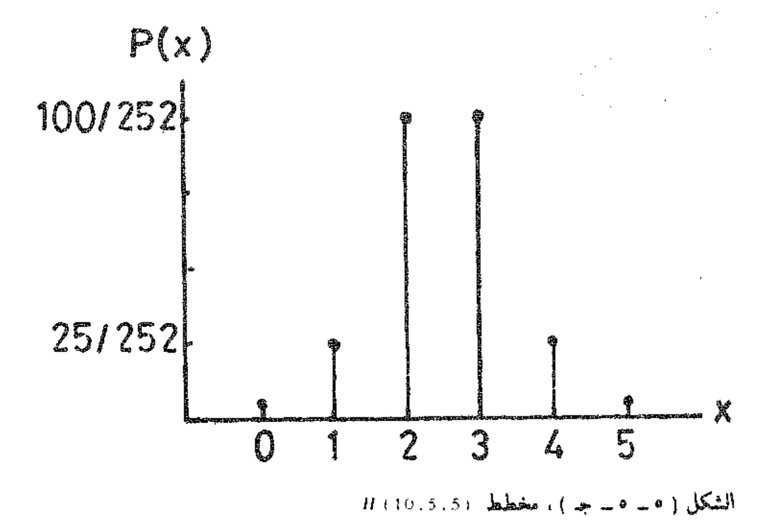


الشكل (٥ ـ ٥ ـ أ)، مخطط H (10.5.3)

max (0 , − 2) ≤ x ≤ min (3 , 5)

الشكل (٥ ـ ٥ ـ ب)، مخطط H (10.5.7)

max (0 , 2) ≤ x ≤ min (5 , 7)



الشكل (٥ ـ ٥ ـ جـ)، مخطط H (10.5.5)

max (0 , 0) ≤ x ≤ min (5 , 5)

ويمكن بيان ان مجموع الكتل الاحتمالية المقترنة بعناصر فضاء **X** مساوٍ للواحد دلالة على كون **P(x)** دالة كتلة احتمالية وعلى النحو الآتي :

$$\sum_{x=0}^{n} P(x ; N, M, n) = \frac{1}{C_n^N} \sum_{x=0}^{n} C_x^M . C_{n-x}^{N-M}$$

لكن وبشكل عام $\sum_{i=0}^{n} C_i^a . C_{n-i}^b = C_n^{a+b}$. عليه فان

$$\sum_{x=0}^{n} C_x^M . C_{n-x}^{N-M} = C_n^N \quad \text{فاذن} \quad \sum_{x=0}^{n} P(x ; N, M, n) = 1$$

٥ ـ ٥ ـ ١: الدالة التوزيعية Distribution function

تعرف الدالة التوزيعية لتوزيع هندسي زائدي على النحو التالي :

$$F(x) = P_r(X \le x) = \sum_{k=c}^{x} \frac{C_x^M . C_{n-x}^{N-M}}{C_n^N}$$

حيث $C = \max(0, n - N + m)$ علماً انه لا يمكن صياغة $F(x)$ بشكل آخر غير الشكل المعرف اعلاه . وقد جرت محاولات عديدة لايجاد صيغ تقريبية (وفق شروط معينة) يمكن من خلالها حساب $F(x)$. هذه الصيغ تستند الى توزيعات متقطعة وأخرى مستمرة وسوف نستعرض بعضاً منها في الفقرات اللاحقة من هذا الفصل . وهنالك جداول خاصة بهذا التوزيع (انظر الجدول ٢ ملحق ب) تبين قيم $P(x)$ عند قيم مختلفة الى N, M, n . فمثلاً عندما $N = 10$, $n = 5, M = 5$ فان

$$\max(0, 0) = 0 \quad , \min(5, 5) = 5$$

وهذا يعني ان $x = 0, 1, 2, 3, 4, 5$. وعندئذٍ فان الكتل الاحتمالية المقترنة بعناصر فضاء X هي :

x :	0	1	2	3	4	5
P(x) :	0·003968	0·099206	0·396825	0·396825	0·099206	0·003968

وبذلك فان التراكم الاحتمالي ، اي $F(x)$. لهذا المثال هو :

x :	0	1	2	3	4	5
F (x) :	0·003968	0·103174	0·499999	0·896824	0·99603	1

واذا كانت $N = 10$, $M = 6$, $n = 6$ فان

$$\max(0, 2) = 2 \ , \min(6, 6) = 6$$

وهذا يعني ان $x = 2, 3, 4, 5, 6$ وعندئذٍ فان الكتل الاحتمالية المقترنة بعناصر X وكذلك التراكم الاحتمالي هي :

x :	2	3	4	5	6
P (x) :	0·071429	0·380952	0·428571	0·114286	0·004762
F (x) :	0·071429	0·452381	0·880952	0·995238	1

٥ ــ ٥ ــ ٢ : الوسط والتباين Mean and Variance

فيما يلي اشتقاق لصيغة الوسط وصيغة التباين للتوزيع الهندسي الزائدي .
وبفرض ان $\max(0, n - N + M) = 0$ وان $\min(n, M) = n$ ، فبالنسبة للوسط فان :

$$\mu_x = \frac{1}{C_n^N} \sum_{x=0}^{n} x\, C_x^M \cdot C_{n-x}^{N-M}$$

$$= \frac{M}{\frac{N}{n}} \sum_{x=1}^{n} \frac{(M-1)!}{(x-1)!\,(M-x)!} \cdot \frac{C_{n-x}^{N-M}}{C_{n-1}^{N-1}}$$

$$= \frac{nM}{N} \sum_{y=0}^{n'} \frac{C_y^{M'} \cdot C_{n'-y}^{N'-M'}}{C_{n'}^{N'}} \qquad , y \sim H(N', M', n')$$

حيث : $n' = n - 1 , M' = M - 1 , N' = N - 1, y = x-1$ فاذن : $\mu_x = \frac{nM}{N}$

اما بالنسبة للتباين فاننا نحتاج لايجاد EX^2 وكما هو مبين بالاتي :

$$EX^2 = E[X(X-1)+X]$$

$$= \sum_{x=0}^{n} x(x-1) . \frac{C_x^M . C_{n-x}^{N-M}}{C_n^N} + \frac{nM}{N}$$

$$= \frac{M(M-1)}{\frac{N(N-1)}{n(n-1)}} \sum_{x=2}^{n} \frac{(M-2)!}{(x-2)!(M-x)!} . \frac{C_{n-x}^{N-M}}{C_{n-2}^{N-2}} + \frac{nM}{N}$$

$$= \frac{nM(n-1)(M-1)}{N(N-1)} \sum_{z=0}^{n'} \frac{C_z^{M'} . C_{n'-z}^{N'-M'}}{C_{n'}^{N'}} + \frac{nM}{N}$$

حيث : $n' = n - 2 , M' = M - 2 , N' = N - 2 , Z = x - 2$ وكأن $Z \sim H(N', M', n')$ فاذن :

$$EX^2 = \frac{nM(n-1)(M-1)}{N(N-1)} + \frac{nM}{N}$$

عليه فان :

$$\sigma_x^2 = EX^2 - (EX)^2$$

$$= \frac{nM(n-1)(M-1)}{N(N-1)} + \frac{nM}{N} - \left(\frac{nM}{N}\right)^2$$

$$= \frac{nM}{N} . \frac{N-M}{N} . \frac{N-n}{N-1} = \mu_x . \frac{(N-M)(N-n)}{N(N-1)}$$

وحيث ان الكسر المضروب بـ μ_x اقل من واحد فذلك يعني ان وسط هدا التوزيع دائماً اكبر من تباينه .

٥ ـ ٥ ـ ٣ : الدالة المولدة للعزوم

ان للتوزيع الهندسي الزائدي دالة مولدة للعزوم ، الا انه من الصعوبة جداً صياغة هذه الدالة بشكل مألوف كالذي لاحظناه في التوزيعات السابقة . وقد امكن التوصل الى صيغة معقدة جداً يصعب التعامل معها تطبيقياً . هذه الصيغة مشتقة بالاعتماد على ما يسمى بـ « الدالة الهندسية الزائدية Hypergeometric function » هذه الصيغة هي :

$$M_X(t) = \frac{(N-n)!(N-M)!}{N!} H(-n; -M; N-M-n+1; e^t)$$

حيث :

$$H(-n; -M; N-M-n+1; e^t) = \sum_{j=0}^{\infty} \frac{(-n)^{[j]}.(-M)^{[j]}(e^t)^j}{(N-M-n+1)^{[j]} j!}$$

وانه بشكل عام ولاي عدد مثل a فان :

$$a^{[j]} = a(a+1)(a+2)\dots(a+j-1)$$

وحيث ان الهدف من الدوال المولدة للعزوم هو توليد عزوم التوزيع لذا يمكن الاستعاضة عنها من خلال ايجاد العزم ذي المرتبة r حول نقطة الاصل الذي يقابل المشتقة ذات المرتبة r للدالة $M_X(t)$ وجعل t مساوية للصفر . فمثلاً لغرض تحديد قيمة العزم الثالث حول الصفر اي EX^3 فان ذلك يمكن ايجاده وفق نفس الاسلوب المتبع في ايجاد الوسط والتباين ، اي ان

$$EX^3 = E[X(X-1)(X-2)] + 3EX(X-1) + EX$$

وان

$$EX^4 = E[X(X-1)(X-2)(X-3)] - 6EX^3 + 11EX^2 - 6EX$$

كذلك وكاسلوب بديل للدالة المولدة للعزوم يمكن ايجاد عزوم هذا التوزيع من خلال ايجاد العزم العاملي ذي المرتبة r ومن خلاله يمكن استنتاج عزوم التوزيع حول نقطة الاصل . وفيما يلي اشتقاق لهذ العزم وسوف نرمز للعزم العاملي ذا المرتبة r بالرمز $\mu_{(r)}$ فاذن

$$\mu_{(r)} = E \prod_{j=1}^{r} (X-j+1)$$

$$= \sum_{x=0}^{n} \prod_{j=1}^{r} (x-j+1) . \frac{C_x^M \; C_{n-x}^{N-M}}{C_n^N}$$

ولغرض السهولة سوف نرمز الى $\prod_{j=1}^{r}(X-j+1)$ بالرمز $X^{(r)}$. عليه فان :

$$\mu_{(r)} = EX^{(r)} = \sum_{x=0}^{n} x^{(r)} . \frac{C_x^M . C_{n-x}^{N-M}}{C_n^N}$$

$$= \sum_{x=r}^{n} x^{(r)} . \frac{C_x^M . C_{n-x}^{N-M}}{C_n^N}$$

$$= \sum_{x=r}^{n} \frac{M!}{(x-r)!(M-x)!} . \frac{C_{n-x}^{N-M}}{C_n^N}$$

$$= \frac{M^{(r)}}{\frac{N^{(r)}}{n^{(r)}}} \sum_{x=r}^{n} \frac{C_{x-r}^{M-r} . C_{n-x}^{N-M}}{C_{n-r}^{N-r}}$$

$$= \frac{n^{(r)} . M^{(r)}}{N^{(r)}} \sum_{y=0}^{n'} \frac{C_y^{M'} . C_{n'-y}^{N'-M'}}{C_{n'}^{N'}}$$

حيث $n' = n - r , M' = M - r , N' = N - r , y = x - r$ وكأن y يتوزع وفق دالة توزيع هندسي زائدي بالمعالم N' , M' , n' . فاذن

$$\mu_{(r)} = \frac{n^{(r)} . M^{(r)}}{N^{(r)}} , r = 1 , 2 , 3 , ...$$

يلاحظ من هذه الصيغة مايلي :

عندما $r = 1$ فان $\mu_{(1)} = \frac{nM}{N} = \mu_x = EX$

عندما $r = 2$ فان $\mu_{(2)} = \frac{n(n-1) . M(M-1)}{N(N-1)} = EX(X-1)$

فاذن

$$EX^2 = \mu_{(2)} - \mu_{(1)}$$

$$= \frac{nM(n-1)(M-1)}{N(N-1)} + \frac{nM}{N}$$

وهي صيغة العزم الثاني التي توصلنا لها عند حسابنا للتباين .

٥ ـ ٥ ـ ٤ : صيغة التراجع Recurrence formula

ان صيغة التراجع في التوزيع الهندسي الزائدي هي :

$$P(x+1) = \frac{(n-x)(M-x)}{(x+1)(N-M-n+x+1)} P(x)$$

ونترك مسألة اشتقاق هذه الصيغة للقاري من خلال قيامه بقسمة $P(x+1)$ على $P(x)$ واجراء بعض الاختزالات بين حدود البسط والمقام ومن ثم التوصل لها .

٥ ـ ٥ ـ ٥ : خاصية التقارب من توزيع ثنائي الحدين
Approximation to the binomial distribution

ان لهذه الخاصية اهمية تطبيقية كبيرة ، حيث انها تسمح باستخدام توزيع ثنائي الحدين كتقريب جيد للتوزيع الهندسي الزائدي عندما يلاحظ ان N عدد كبير (نظرياً $N \to \infty$) وان $\frac{M}{N}$ يستقر نحو عدد ثابت مثل P . ان المطلوب برهنته هنا هو ان

$$\lim_{N \to \infty} H(N, M, n) \to b\left(n, p = \frac{M}{N}\right)$$

البرهان :ان

$$P(x; N, M, n) = \frac{C_x^M \cdot C_{n-x}^{N-M}}{C_n^N}$$

$$= \frac{M!}{x!(M-x)!} \cdot \frac{(N-M)!}{(n-x)!.(N-M-n+x)!} \cdot \frac{n!(N-n)!}{N!}$$

$$= \frac{n!}{x!(n-x)!} \cdot \frac{M!}{(M-x)!} \cdot \frac{(N-M)!}{(N-M-n+x)!} \cdot \frac{(N-n)!}{N!}$$

ان

$$\frac{n!}{x!(n-x)!} = C_x^n$$

وان

$$\frac{M!}{(M-x)!} = \frac{M(M-1)(M-2)\dots(M-x+1)(M-x)!}{(M-x)!}$$

$$= M(M-1)(M-2)\dots(M-x+1)$$

مع ملاحظة ان عدد الحدود المضروبة ببعضها في الصيغة الاخيرة هو (x) حد .
كذلك فان :

$$\frac{(N-M)!}{(N-M-n+x)!} = \frac{(N-M).(N-M-1)\dots(N-M-n+x+1)(N-M-n+x)!}{(N-M-n+x)!}$$

$$= (N-M)(N-M-1)\dots(N-M-n+x+1)$$

مع ملاحظة ان عدد الحدود المضروبة ببعضها هنا هو (n − x) حد .
واخيراً فان :

$$\frac{(N-n)!}{N!} = \frac{(N-n)!}{N(N-1)(N-2)\dots(N-n+1)(N-n)!}$$

$$= \frac{1}{N(N-1)(N-2)\dots(N-n+1)}$$

وهنا نلاحظ ان عدد الحدود المضروبة ببعضها في مقام الكسر هو (n) حد .
فاذن :

$$p(x;N,M,n) = C^n_x \cdot \frac{M(M-1)\ldots(M-x+1).(N-M)(N-M-1)\ldots(N-M-n+x+1)}{N(N-1)(N-2)\ldots(N-n+1)}$$

وهنا نلاحظ ان عدد الحدود المضروبة ببعضها في البسط مساوٍ لعدد الحدود المضروبة ببعضها في المقام البالغة n حد . عليه وبقسمة البسط والمقام على N^n نحصل على :

$$P(x) = C^n_x \cdot \frac{\frac{M}{N}\left(\frac{M}{N}-\frac{1}{N}\right)\ldots\left(\frac{M}{N}-\frac{x-1}{N}\right).\left(1-\frac{M}{N}\right)\left(1-\frac{M}{N}-\frac{1}{N}\right)\ldots\left(1-\frac{M}{N}-\frac{n-x-1}{N}\right)}{\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\ldots\left(1-\frac{n-1}{N}\right)}$$

الان بفرض ان $M \leq N$ ، $P = \frac{M}{N}$ وجعل $N \to \infty$ نحصل على :

$$\lim_{N\to\infty} P(x;N,M,n) = C^n_x . P.P.\ldots P.(1-P)(1-P)\ldots.(1-P)$$

$$= C^n_x P^x . (1-P)^{n-x}, \quad = 0, 1, 2, \ldots, n$$

وعلى ضوء ماتقدم فان

$$\mu_x = \frac{nM}{N} = nP$$

وان

$$\lim_{N\to\infty} \sigma^2_x = \lim_{N\to\infty} \frac{nM}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1}$$

$$= nP(1-P) \lim_{N\to\infty} \left(\frac{N}{N-1} - \frac{n}{N-1}\right) = nP(1-P)$$

وكما سبق وان ذكرنا فان خاصية التقارب من توزيع ثنائي الحدين ذات اهمية تطبيقية كبيرة حيث انها تسمح باستخدام جداول توزيع ثنائي الحدين (عندما N كبيرة) كبديل لجداول التوزيع الهندسي الزائدي عندما يتطلب الامر حساب احتمال معين . فمثلاً اذا كان $X \sim H(100, 40, 15)$ وتطلب الامر حساب الكتل الاحتمالية المقترنة بعناصر فضاء X اي $x = 0, 1, 2, \ldots, 15$ فاننا نجابه في

هذه الحالة صعوبة كبيرة في حساب هذه الكتل باستخدام التوزيع الهندسي الزائدي (وخصوصاً اذا ماعلمنا ان جداول هذا التوزيع تكون وفي اغلب الاحوال غير معرفة عند قيم كبيرة الى N حيث ان جدول هذا التوزيع الموجود في الملحق ب ينتهي بقيمة $N = 10$) في حيث يمكن حسابها وبسهولة باستخدام توزيع ثنائي الحديد طالما ان $N = 100$ عدد كبيرنسبياً وعلى النحو الآتي : ان

$$P = \frac{M}{N} = \frac{40}{100} = 0{\cdot}4 \quad \therefore q = 1 - P = 0{\cdot}6$$

فاذن $X \sim b(15, 0{\cdot}4)$. ومن جداول توزيع ثنائي الحدين (او باستخدام اية طريقة اخرى لحساب $P(x)$ نلاحظ ان :

$$P(0) = 0{\cdot}0005,\ P(1) = 0{\cdot}0047,\ P(2) = 0{\cdot}0219,\ \ldots,\ P(15) \simeq 0.$$

كذلك فان هذه الخاصية تسمح لنا بايجاد عزوم التوزيع الهندسي الزائدي باستخدام عزوم توزيع ثنائي الحدين . فمثلاً

$$\mu_x = nP = 6,\ \sigma_x^2 = nPq = 3{\cdot}6$$

٥ ـ ٥ ـ ٦ : امثلة

مثال (١) : اذا كان $X \sim H(10, 6, 4)$ فان :

$$1 - \ P(x) = \frac{C_x^6 \cdot C_{4-x}^4}{C_4^{10}}\ ; x = 0, 1, 2, 3, 4$$

$$2 - \ \mu_x = \frac{nM}{N} = 2{\cdot}4,\ \sigma_x^2 = \mu_x \cdot \frac{(N-M)(N-n)}{N(N-1)} = 0{\cdot}64$$

$$3 - \ P_r(X = 2) = 0{\cdot}4286.$$

مثال (٢) : لجدول التوزيع التكراري الاتي يطلب توفيق توزيع هندسي :

x :	0	1	2	3	4	5
f_x :	2	100	410	390	95	3

الحل :

ان المعلوم فقط عن معالم التوزيع الهندسي الزائدي في هذا المثال هو ان $n = 5$ والمطلوب البحث عن قيمة N, M وعلى النحو الآتي :

ان

$$\bar{x} = \frac{\sum_{x=0}^{5} x f_x}{\sum_{x=0}^{5} f_x} = \frac{2485}{1000} = 2.485$$

وان

$$S_x^2 = \frac{\sum_{x=0}^{5} x^2 f_x}{\sum_{x=0}^{5} f_x} - \bar{x}^2 = \frac{6845}{1000} - (2.485)^2 = 0.67$$

الان بجعل :

$$\bar{x} = \frac{nM}{N} = \frac{5M}{N} \rightarrow 2.485 = \frac{5M}{N}$$

$$\therefore M = 0.497\,N$$

وان

$$S_x^2 = \bar{x}\,.\,\frac{(N-M)(N-n)}{N(N-1)} = (2.485).\,\frac{(N-0.497N)(N-5)}{N(N-1)} = 0.67$$

وبحل الصيغة الاخيرة نسبة الى N نلاحظ ان $N \simeq 10$. فاذن $M = 0.497\,N \simeq 5$ ونفرض ان $X \sim H(10, 5, 5)$. عليه يمكن الرجوع لجدول التوزيع الهندسي الزائدي (جدول ٢ ملحق ب) لايجاد الكتل الاحتمالية المقترنة

بعناصر فضاء هذا المتغير والتي على اساسها يتم حساب التكرارات المتوقعة من خلال $E.f = P(x)\,\Sigma f_x$ والمبينة في الجدول الاتي :

قيم X : (x)	0	1	2	3	4	5
التكرار المشاهد :	2	100	410	390	95	3
التكرار المتوقع :	4	99	397	397	99	4

مثال (٣) : صندوق يحتوي على 20 كرة . 12 كرة منها حمراء والبقية سوداء اختيرت عينة عشوائية من هذا الصندوق قوامها 8 كرات . ماهو احتمال :

أ ـ الحصول على ثلاث كرات حمراء ضمن هذه العينة .

ب ـ الحصول على كرتان حمراوان على الاقل .

الحل :

افرض ان X يشير الى عدد الكرات الحمراء المسحوبة ضمن العينة .

واضح هنا ان $X \sim HG(20, 12, 8)$. وهذا يعني ان :

$$P(x) = \frac{C_x^{12} \cdot C_{8-x}^{8}}{C_8^{20}} \;;\; 0 \leq x \leq 8$$

عليه :

أ ـ احتمال الحصول على ثلاث كرات حمراء ضمن هذه العينة هو :

$$P_r(X = 3) = \frac{C_3^{12} \cdot C_5^{8}}{C_8^{20}} = 0{\cdot}097801$$

ب ـ احتمال الحصول على كرتان حمراوان على الاقل هو :

$$P_r(X \geq 2) = 1 - P_r(X < 2) = 1 - P_r(X \leq 1)$$

لكن

$$P_r(X \leq 1) = P_r(X = 0) + P_r(X = 1)$$

$$= \frac{C_0^{12} \cdot C_8^{8}}{C_8^{20}} + \frac{C_1^{12} \cdot C_7^{8}}{C_8^{20}} = 0{\cdot}0008$$

فاذن

$$P_r(X \geq 2) = 1 - 0{\cdot}0008 = 0{\cdot}999\,2$$

تمارين عن التوزيع الهندسي الزائدي

٥ ـ ١٩ : إذا علمت ان $X \sim H(9, 4, 5)$. جد مايلي :

أ ـ دالة الكتلة الاحتمالية للمتغير X مع رسم مخطط هذه الدالة .

ب ـ الوسط والتباين لهذا التوزيع .

جـ ـ العزم الثالث والرابع حول نقطة الاصل .

د ـ العزم العاملي الثالث والرابع .

هـ ـ $P_r(X \geq 1), P_r(X \leq 3)$.

٥ ـ ٢٠ : افرض ان $X \sim H(50, 35, 10)$ جد $P_r(X \leq 5), P_r(X \geq 2)$

٥ ـ ٢١ : اذا علمت ان $X_1 \sim H(50, 40, 10)$ مستقل عن $X_2 \sim H(75, 60, 12)$ وافرض ان $Y = X_1 + X_2$. جد :

أ ـ $P_r(Y \leq 5), P_r(Y \geq 3)$.

ب ـ الدالة المولدة لعزوم المتغير Y .

٥ ـ ٢٢ : ليكن $X \sim H(N, M, n)$ وليكن $\mu_{(r)}$ يمثل العزم العاملي ذا المرتبة r . برهن ان :

$$\lim_{N \to \infty} \mu_{(r)} = P^r . n^{(r)}, P = \frac{M}{N}$$

٥ ـ ٢٣ : افرض ان $X_1 \sim b(n_1, p)$ مستقل عن $X_2 \sim b(n_2, p)$. برهن ان التوزيع الشرطي للمتغير X_1 علماً ان $x_1 + x_2 = M$ هو توزيع هندسي زائدي دالته الاحتمالية هي :

$$P(x_1 = K \mid x_1 + x_2 = M) = \frac{C_k^{n_1} . C_{M-k}^{n_2}}{C_M^{n_1+n_2}}$$

$$\max(0, M - n_2) \leq x_1 \leq \min(n_1, M)$$

٥ ـ ٦ : توزيع پواسون Poisson distribution

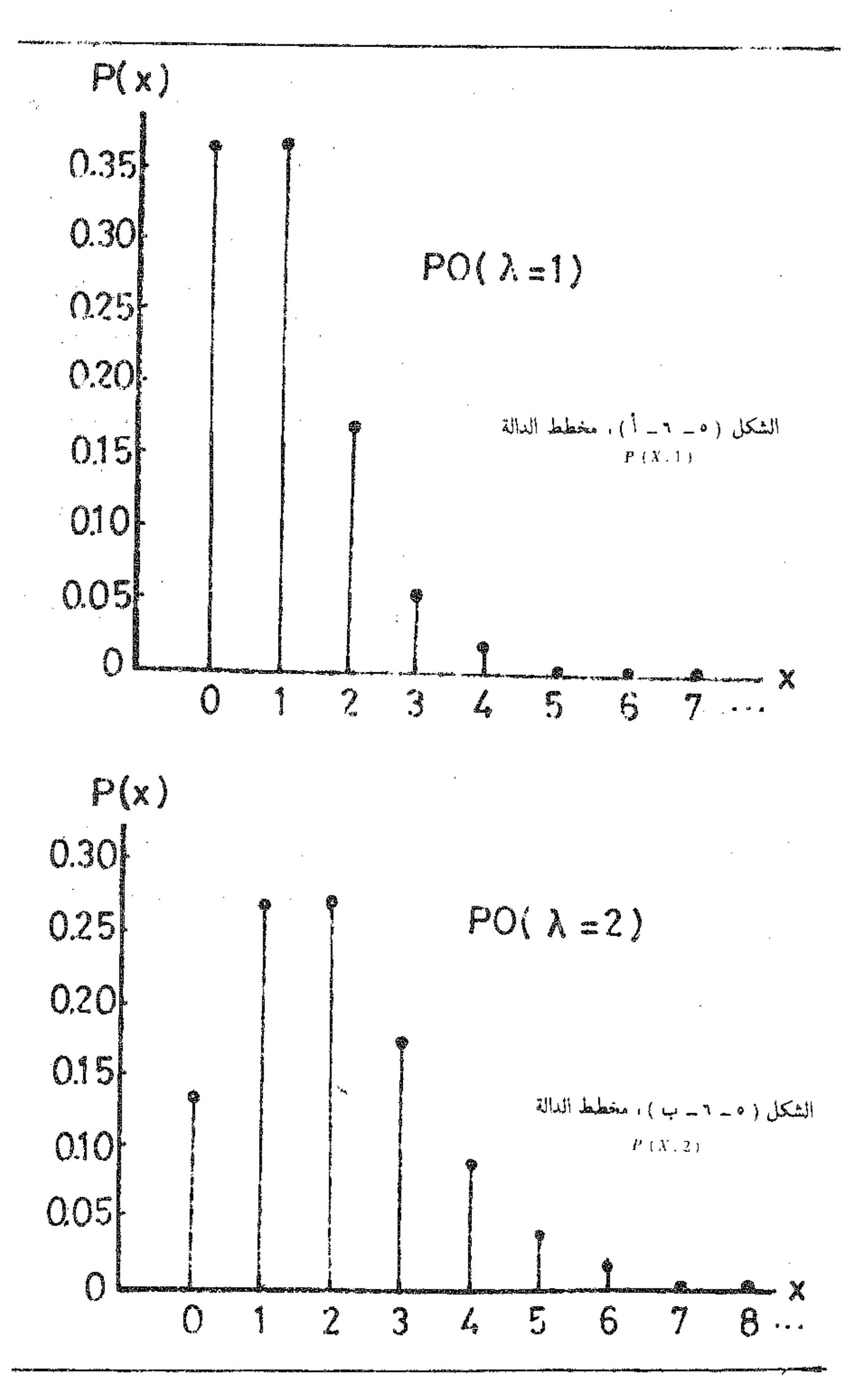
يعد توزيع پواسون احد التوزيعات المتقطعة المهمة جداً في الكثير من التطبيقات الاحصائية . ويسمى في بعض الاحيان توزيع الحوادث النادرة الوقوع كحوادث سقوط الطائرات ، عدد النداءات الهاتفية المستلمة من قبل بدالة هاتف خلال فترة زمنية محددة ، عدد الوحدات المعيبة في انتاج واسع لمصنع معين وغيرها من الامثلة التي تتصف بطابع الندرة . ان اول من اشتق هذا التوزيع هو العالم الرياضي ـ الفيزيائي الفرنسي Simeon Denis Poisson (١٧٨١ ـ ١٨٤٠) الذي تمكن من اشتقاق هذا التوزيع كحالة تقاربية من توزيع ثنائي الحدين ونشر اشتقاقه هذا عام ١٨٣٧ مطلقاً اسمه على هذا التوزيع . وفيما يلي تعريف هذا التوزيع .

يقال ان المتغير العشوائي X هو ذو توزيع پواسون اذا كانت دالة الكتلة الاحتمالية لهذا المتغير تأخذ الشكل التالي :

$$P(x,\lambda) = \frac{\lambda^{x} \cdot e^{-\lambda}}{x!} \; ; x = 0, 1, 2, \ldots$$

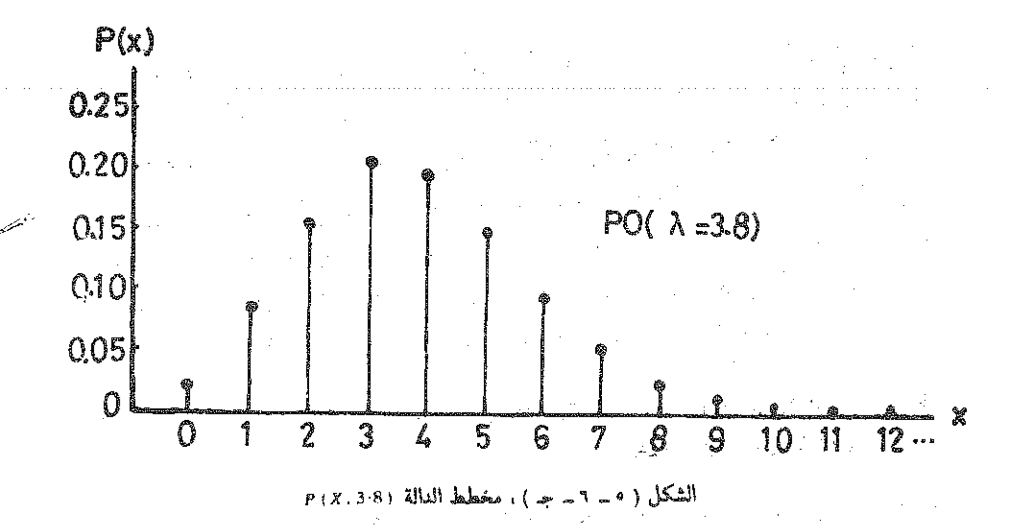
$$= 0 \qquad \text{other wise}$$

حيث $\lambda > 0$ تمثل معلمه هذا التوزيع . وبالرموز فان $X \sim P0(\lambda)$ والاشكال (٥ ـ ٦) توضح مخطط دالة هذا التوزيع .



الشكل (٥ ـ ٦ ـ أ) ، مخطط الدالة
$P(X, 1)$

الشكل (٥ ـ ٦ ـ ب) ، مخطط الدالة
$P(X, 2)$



الشكل (٥ ـ ٦ ـ جـ) ، مخطط الدالة $P(X, 3.8)$

ويمكن بيان ان مجموع الكتل الاحتمالية المقترنة بعناصر فضاء X مساوٍ للواحد وكما يلي :

$$\sum_{x=0}^{\infty} P(x;\lambda) = \sum_{x=0}^{\infty} \frac{\lambda^x . e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

لكن $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$ يعبر عن مجموع حدود سلسلة لانهائية تتقارب من e^{λ} حسب سلسلة تايلر . عليه فان

$$\sum_{x=0}^{\infty} P(x;\lambda) = e^{-\lambda} . e^{\lambda} = 1$$

٥ ـ ٦ ـ ١ : الدالة التوزيعية

ان الدالة التوزيعية في توزيع پواسون بشكل عام معطاة وفق الآتي :

$$F(x) = P_r(X \leq x) = e^{-\lambda} \sum_{k=0}^{x} \frac{\lambda^k}{K!}$$

علماً انه لايمكن صياغة هذه الدالة بشكل آخر غير الشكل الموضح اعلاه . ان مسألة التعامل مع هذه الدالة وبهذا الشكل تطبيقياً تبدو معقدة بعض الشيء وخصوصاً عند حساب $F(x)$ لقيم كبيرة الى X . الا ان ذلك يمكن جعله امراً سهلاً في حالة

برمجة هذه الدالة باحدى لغات البرمجة المعروفة على حاسب الكتروني من شأنه حساب هذه الدالة لاية قيمة معطاة الى X مثل x ولاية قيمة مخصصة الى المعلمة λ على اية حال تم اقتراح العديد من الصيغ التقريبية للدالة $F(x)$ كبدائل للصيغة اعلاه بعضاً منها استند الى التوزيع الطبيعي المعياري (لاحظ الفقرة ٦ ـ ٢) والبعض الاخر استند الى توزيع مربع كاي (لاحظ الفقرة ٨ ـ ٣ ـ ١) . وهنالك جداول خاصة بهذا التوزيع (لاحظ الجدول ٣ ملحق ب) تبين الكتل الاحتمالية المقترنة بعناصر فضاء X عند قيم مختلفة للمعلمة λ . فمثلاً عندما $\lambda = 4$ فان :

$$P(0) = 0{\cdot}0183 , P(1) = 0{\cdot}0733 , P(2) = 0{\cdot}1465 , \ldots P(14) = 0{\cdot}0001 , P(15) \simeq 0$$

كذلك يمكن تعريف $F(x)$ في توزيع پواسون باستخدام ما يسمى بـ « تكامل كاما الناقص "Incomplete gamma integral » المعطاة صيغته بما يلي :

$$I_x = \frac{1}{x!} \int_{\lambda}^{\infty} e^{-t} . t^x \, dt \; ; x = 0, 1, 2 \ldots$$

البرهان : باستخدام طريقة التكامل بالتجزئة من خلال الفرض ان

وان
$$u = t^x \rightarrow du = xt^{x-1}$$
$$dv = e^{-t}dt \rightarrow V = -e^{-t}$$

فاذن

$$I_x = \frac{1}{x!} \left[-e^{-t}t^x + x \int e^{-t}t^{x-1}dt \right]_{\lambda}^{\infty}$$

$$= \frac{\lambda^x e^{-\lambda}}{x!} + \frac{1}{(x-1)!} \int_{\lambda}^{\infty} e^{-t} . t^{x-1}dt$$

$$= P(X = x ; \lambda) + I_{x-1}$$

وباستخدام التكامل بالتجزئة مرة اخرى لحل I_{x-1} بنفس الاجراء الموضح اعلاه نحصل على :

$$I_{x-1} = \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} + I_{x-2}$$

فاذن

$$I_x = P(X = x ; \lambda) + P(X = x - 1 ; \lambda) + I_{x-2}$$

ولو استمر الحال باستخدام التكامل بالتجزئة فاننا سوف نحصل على :

$$I_x = \frac{\lambda^x e^{-\lambda}}{x!} + \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} + \frac{\lambda^{x-2} e^{-\lambda}}{(x-2)!} + \dots + \frac{\lambda e^{-\lambda}}{1!} + e^{-\lambda}$$

$$= \sum_{k=0}^{x} \frac{\lambda^k e^{-\lambda}}{k!} = F(x)$$

٥ ــ ٦ ــ ٢ : الوسط والتباين

ان الخاصية المميزة لتوزيع پواسون عن بقية التوزيعات المتقطعة الاخرى هي ان متوسط هذا التوزيع مساوٍ لتباينه ويكون مساوياً لقيمة المعلمة λ.

البرهان :

$$\mu_x = e^{-\lambda} \sum_{x=0}^{\infty} x \cdot \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} x \cdot \frac{\lambda^x}{x!}$$

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} \; ; y = x - 1$$

لكن حسب متسلسلة تايلر فان المجموع الاخير يتقارب من e^{λ} . عليه فان :

$$\mu_x = \lambda e^{-\lambda} . e^{\lambda} = \lambda$$

كذلك فان

$$\sigma_x^2 = EX^2 - (EX)^2$$

لكن

$$EX^2 = E[X(X-1)+X] = e^{-\lambda} \sum_{x=0}^{\infty} x(x-1)\frac{\lambda^x}{x!} + \lambda$$

$$= \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda$$

$$= \lambda^2 e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} + \lambda \quad ; z = x-2$$

$$= \lambda^2 e^{-\lambda} . e^{\lambda} + \lambda \qquad \therefore EX^2 = \lambda^2 + \lambda$$

فاذن

$$\sigma_x^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

وهذا يعني ان

$$\mu_x = \sigma_x^2 = \lambda$$

٥ ـ ٦ ـ ٣ : الدالة المولدة للعزوم

يمتلك توزيع پواسون دالة مولدة لعزومه حول نقطة الاصل . هذه الدالة هي :

$$M_X(t) = e^{\lambda(e^t - 1)}$$

البرهان :

$$M_X(t) = e^{-\lambda} \sum_{x=0}^{\infty} e^{tx} . \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

وحسب سلسلة تايلر فان المجموع الاخير متقارب من $e^{\lambda e^t}$. فاذن

$$M_X(t) = e^{-\lambda} . e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

كذلك ووفق نفس الاجراء يمكن بيان ان الدالة المميزة لتوزيع پواسون هي :

$$\phi(t) = e^{\lambda(e^{it} - 1)}$$

وان

$$K_X(t) = \ln M_X(t) = \lambda(e^t - 1)$$

وواضح بان :

$$K_X'(t) = \lambda e^t \rightarrow K_X'(0) = \mu_x = \lambda$$

وان

$$K_X''(t) = \lambda e^t \rightarrow K_X''(0) = \sigma_x^2 = \lambda$$

وبشكل عام فان

$$K_X^{(r)}(t) = \lambda e^t \rightarrow K_X^{(r)}(0) = \lambda$$

٥ ـ ٦ ـ ٤ : صيغة التراجع Recurrence formula

ان صيغة التراجع في توزيع پواسون هي :

$$P(x+1) = \frac{\lambda}{x+1} . P(x)$$

البرهان :

$$\frac{P(x+1)}{P(x)} = \frac{\dfrac{\lambda^{x+1}.e^{-\lambda}}{(x+1)!}}{\dfrac{\lambda^x.e^{-\lambda}}{x!}} = \frac{\lambda}{x+1}$$

فاذن

$$P(x+1) = \frac{\lambda}{x+1} . P(x)$$

وعن طريق هذه الصيغة يمكن تحديد الكتل الاحتمالية المقترنة بعناصر فضاء X دون اللجوء للتعويض في الدالة $P(x)$ وانما يتطلب ذلك فقط حساب $P(0)$ ومن ثم تتحدد بقية الكتل الاحتمالية اللاحقة للعنصر $X = 0$ من خلال هذه الصيغة وكما هو موضح بالآتي : ان $P(0) = e^{-\lambda}$ فاذن :

$$P(1) = \lambda P(0), P(2) = \frac{\lambda}{2} P(1), P(3) = \frac{\lambda}{3} P(2), \ldots$$

وبشكل عام فان :

$$P(j) = \frac{\lambda}{j} P(j-1), j = 1, 2, 3, \ldots$$

٥ ـ ٦ ـ ٥ : خاصية الجمع في توزيع پواسون :

افرض ان $X_1 , X_2 , \ldots , X_n$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim P0(\lambda_i)$ عندئذٍ فان $Y = \sum_{i=1}^{n} X_i \sim P0\left(\sum_{i=1}^{n} \lambda_i\right)$

البرهان : لتكن $M_Y(t)$ تمثل الدالة المولدة لعزوم Y حول نقطة الاصل عندئذٍ

$$M_Y(t) = Ee^{tY} = E \prod_{i=1}^{n} e^{tX_i}$$

وحيث ان هذه المتغيرات مستقلة تصادفياً فذلك يعني ان

$$M_Y(t) = \prod_{i=1}^{n} Ee^{tX_i} = \prod_{i=1}^{n} M_{X_i}(t)$$

وحيث ان $X_i \sim P0(\lambda_i)$. فإذن $M_{X_i}(t) = e^{\lambda_i(e^t-1)}$ عليه فان :

$$M_Y(t) = \prod_{i=1}^{n} e^{\lambda_i(e^t-1)}$$

$$= e^{\sum_{i=1}^{n} \lambda_i(e^t-1)}$$

ويلاحظ ان الصيغة الاخيرة تمثل الدالة المولدة لعزوم متغير ذي توزيع پواسون بالمعلمة $\sum_{i=1}^{n} \lambda_i$ فاذن $Y = \sum_{i=1}^{n} X_i \sim P0\left(\sum_{i=1}^{n} \lambda_i\right)$

٥ ـ ٦ ـ ٦ : توزيع پواسون كحالة تقاربية من توزيع ثنائي الحدين .

سبق وان اشرنا في بداية الفقرة (٥ ـ ٦) ان توزيع پواسون هو توزيع مشتق من توزيع ثنائي الحدين كحالة تقاربية . وفيما يلي برهان لذلك :

البرهان :

ان المطلوب برهنته هنا هو :

$$\lim_{n \to \infty} b(n, p) \to P0(\lambda = np)$$

ليكن $X \sim b(n, P)$ فاذن

$$P(x : n, p) = C^n_x P^x (1 - P)^{n-x} ; x = 0, 1, 2, \ldots, n$$

$$= \frac{n!}{x!(n-x)!} \left(\frac{P}{1-P} \right)^x . (1 - P)^n$$

وكما هو معلوم فان الوسط في توزيع ثنائي الحدين هو nP . ولنفرض ان $\lambda = nP$ فاذن $P = \frac{\lambda}{n}$ وان

$$P(x; n, p) = \frac{n(n-1)(n-2)\ldots(n-x+1)}{x!} . \left(\frac{\lambda/n}{1 - \lambda/n} \right)^x . \left(1 - \frac{\lambda}{n} \right)^n$$

$$= \frac{\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{x-1}{n}\right)}{x!} . \frac{\lambda^x}{\left(1 - \frac{\lambda}{n}\right)^x} . \left(1 - \frac{\lambda}{n}\right)^n$$

الان بجعل $n \to \infty$ فان $\lim_{n \to \infty} \frac{\lambda}{n} \to 0$ إن :

$$\lim_{n \to \infty} \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{x-1}{n}\right) \to 1$$

وان

$$\lim_{n \to \infty} \left(1 - \frac{\lambda}{n}\right)^x \to 1$$

كذلك فان :

$$\lim_{n \to \infty} \left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$$

ونترك للقارىء برهنة الحالة الاخيرة من خلال فك المقدار $\left(1-\frac{\lambda}{n}\right)^{n}$ باستخدام نظرية ثنائي الحدين ومن ثم جعل $n \to \infty$. عليه فان .

$$\lim_{n \to \infty} b(n, p) = \frac{\lambda^{x} . e^{-\lambda}}{x!} \; ; x = 0, 1, 2, \ldots$$

فاذن نستنتج ان $X \sim P0(\lambda = np)$.

ان لهذه الخاصية اهمية تطبيقية كبيره . حيث انها تسمح باستخدام توزيع پواسون كبديل لتوزيع ثنائي الحدين عند حساب احتمال معين بمجرد ملاحظتنا ان n كبيرة وان احتمال نجاح المحاولة p عدد صغير . فمثلاً اذا كان $X \sim b(100, 0{\cdot}04)$ وتطلب الامر حساب الاحتمال $P_r(X = 10)$ فان قيمة هذا الاحتمال باستخدام توزيع ثنائي الحدين هو :

$$P_r(X = 10) = C_{10}^{100}(0{\cdot}04)^{10} . (0{\cdot}96)^{90} = 0{\cdot}0047 \simeq 0{\cdot}005$$

لاحظ ان هنالك بعض الصعوبة في حساب هذا الاحتمال . لكن لو تم استخدام توزيع پواسون فان قيمة هذا الاحتمال هي :

$$P_r(X = 10) = \frac{\lambda^{x} e^{-\lambda}}{x!} = \frac{4^{10} e^{-4}}{10!} = 0{\cdot}0053 \simeq 0{\cdot}005 \; , \lambda = np = 4$$

علماً انه كلما كانت n كبيرة و p صغير فان مقدار الفرق المطلق ما بين هذين الاحتمالين يتضاءل مقترباً من الصفر .

٥ ـ ٦ ـ ٧ : توزيع پواسون كحالة تقاربية من توزيع ثنائي الحدين السالب .

يمكن التوصل ايضاً الى امكانية اشتقاق توزيع پواسون كحالة تقاربية من توزيع ثنائي الحدين السالب . فاذا كان $X \sim Nb(r, P)$ عندئذٍ فان

$$\lim_{r \to \infty} Nb(r, p) \to P0\left(\lambda = \frac{rq}{p}\right)$$

البرهان : حيث ان $X \sim Nb(r, P)$ عندئذٍ :

$$P(x; r, P) = C_x^{x+r-1} \cdot p^r \cdot q^x ; x = 0, 1, 2, \ldots$$

$$= \frac{(x+r-1)!}{(r-1)!\,x!} p^r \cdot q^x$$

$$= \frac{(x+r-1)(x+r-2)\ldots(r+1)(r)}{x!} p^r \cdot q^x$$

الان بفرض ان : $P = \frac{1}{Q}$, $q = \frac{S}{Q}$ حيث S, Q عددان حقيقيان وان $Q - S = 1$ طالما ان $p + q = 1$ فاذن :

$$P(x; r, P) = \frac{(x+r-1)(x+r-2)\ldots(r+1)(r)}{x!} \cdot \left(\frac{1}{Q}\right)^r \left(\frac{S}{Q}\right)^x$$

واضح ان عدد الحدود المضروبة مع بعضها في بسط الكسر الاول من الصيغة السابقة هو x ، فاذن وبضرب هذا الكسر بـ r^x وقسمته على نفس المقدار وجعل $Q = 1 + S$ نحصل على :

$$P(x, r, P) = \frac{\left(1 + \frac{x-1}{r}\right)\left(1 + \frac{x-2}{r}\right)\ldots\left(1 + \frac{1}{r}\right)(1)}{x!} .$$

$$r^x \left(\frac{1}{1+S}\right)^r \cdot \left(\frac{S}{1+S}\right)^x$$

ويجعل متوسط توزيع پواسون λ مساوٍ لمتوسط توزيع ثنائي الحدين السالب اي $\frac{rq}{P} = \lambda$ فذلك يعني ان :

$$\lambda = rq \cdot Q = r.S \rightarrow S = \frac{\lambda}{r}$$

فاذن :

$$P(x;r,P) = \frac{\lambda^x}{x!}\left(1+\frac{x-1}{r}\right)\left(1+\frac{x-2}{r}\right)\cdots$$

$$\left(1+\frac{1}{r}\right)\cdot\left(1+\frac{\lambda}{r}\right)^{-r}\left(1+\frac{\lambda}{r}\right)^{-x}$$

وبجعل $r \to \infty$ نلاحظ ان :

$$\lim_{r\to\infty} P(x;r,P) = \frac{\lambda^x}{x!}(1).(1)\ldots(1).e^{-\lambda}.(1)$$

$$= \frac{\lambda^x e^{-\lambda}}{x!} \; ; x = 0, 1, 2, \ldots$$

ونترك للقارىء برهنة ان $\lim_{r\to\infty}\left(1+\frac{\lambda}{r}\right)^{-r} = e^{-\lambda}$ من خلال فك المقدار $\left(1+\frac{\lambda}{r}\right)^{-r}$ باستخدام نظرية ثنائي الحدين ومن ثم جعل $r \to \infty$

وتكمن الاهمية التطبيقية لهذه الخاصية في امكانية استخدام توزيع پواسون كبديل لتوزيع ثنائي الحدين السالب في حالة حساب احتمال معين عند ملاحظتنا ان المعلمة r عدد كبير وان P قريبة من الواحد. فمثلاً اذا كان $X \sim Nb(200, 0{\cdot}99)$ وتطلب الامر حساب $P_r(X=5)$ فان ذلك يتم وفق الاتي :

$$\lambda = \frac{rq}{P} = \frac{(200)(0{\cdot}01)}{0{\cdot}99} = 2{\cdot}02$$

$$\therefore P_r(X=5) = \frac{(2{\cdot}02)^5 e^{-2\cdot 02}}{5!} = 0{\cdot}0371792$$

لاحظ السهولة في حساب قيمة هذا الاحتمال. في حين لو تم استخدام توزيع ثنائي السالب فان

$$P_r(X = 5) = C^{204}_{199}(0{\cdot}99)^{200}.(0{\cdot}01)^5 = 0{\cdot}0375457$$

وهنا توجد بعض الصعوبه في حسابه وان مقدار الفرق البالغ 0.0003665 ناتج بسبب ان قيمة r ليست كبيرة جداً وان هذا الفرق يتضاءل كلما كبرت قيمة r

٥ ــ ٦ ــ ٨ : توزيع پواسون كحالة تقاربية من التوزيع الهندسي الزائدي .

يمكن استتاج خاصية اخرى لتوزيع پواسون وهي امكانية استنتاجه كحالة تقاربية من التوزيع الهندسي الزائدي . ان عملية الاستنتاج سوف لاتتم بطريقة البرهان وانما من خلال دراسة علاقة توزيع ثنائي الحدين بالتوزيع الهندسي الزائدي (لاحظ الفقرة ٥ ــ ٥ ــ ٥) وعلاقة توزيع پواسون بتوزيع ثنائي الحدين (لاحظ الفقرة (٥ ــ ٦ ــ ٦) وكما يلي :

بفرض ان $X \sim H(N,M,n)$. لاحظنا في الفقرة (٥ ــ ٥ ــ ٥) ان :

$$\lim_{N\to\infty} H(N,M,n) \to b(n,P), \quad P = \frac{M}{N}$$

كذلك لاحظنا في الفقرة (٥ ــ ٦ ــ ٦) ان

$$\lim_{n\to\infty} b(n,P) \to P0(\lambda); \lambda = nP$$

فاذن نستنتج ان

$$\lim_{\substack{N\to\infty \\ n\to\infty}} H(N,M,n) \to P0(\lambda); \lambda = \frac{nM}{N}$$

فمثلاً لو كان $X \sim H(200, 150, 50)$ وتطلب الامر حساب $P_r(X = 30)$ نلاحظ وجود صعوبة كبيرة في حساب هذا الاحتمال باستخدام التوزيع الهندسي الزائدي في حين يمكن الحصول على قيمة تقريبية له من خلال توزيع پواسون وكما يلي :

فاذن $\lambda = \frac{nM}{N} = 37{\cdot}5$

$$P_r(X = 30) = \frac{(37{\cdot}5)^{30}.e^{-37{\cdot}5}}{30!} = 0{\cdot}0324514$$

$$\simeq C^{150}_{30}.C^{50}_{20}/C^{200}_{50}$$

٥ ـ ٦ ـ ٩ : امثلة

مثال (١) : اذان كان $X \sim P0(5)$ عندئذٍ

$$1-\ P(x;5) = \frac{5^x e^{-5}}{x!}, x = 0, 1, 2, \ldots$$

$$2-\ \mu_x = \sigma_x^2 = 5$$

$$3-\ M_X(t) = e^{5(e^t-1)},\ K_X(t) = 5(e^t - 1)$$

$$4-\ P_r(X=0) = e^{-5},\ P_r(X \geq 1) = 1 - P_r(X=0) = 1 - e^{-5}$$

مثال (٢) : اذا كان $X_1 \sim P0(4)$ مستقل عن $X_2 \sim P0(6)$ وان $Y = X_1 + X_2$ عندئذٍ

$$1-\ Y \sim P0(10), P(y) = \frac{10^y . e^{-10}}{y!}, y = 0, 1, 2, \ldots$$

$$2-\ \mu_y = \sigma_y^2 = 10$$

$$3-\ P_r(Y \leq 1) = P(0) + P(1) = 11e^{-10}$$

مثال (٣) : لجدول التوزيع التكراري الاتي يطلب توفيق توزيع پواسون .

x :	0	1	2	3	4	5	6	7
f_x :	50	160	130	90	66	4	0	0

الحل :

لغرض التوصل الى قيمة λ نجعل الوسط الحسابي لهذا التوزيع ، اي $\bar{x}$ ، مساوٍ الى λ . اي :

فاذن

$$\bar{x} = \frac{\sum_{x=0}^{7} x f_x}{\sum_{x=0}^{7} f_x} = \frac{974}{500} = 1{\cdot}948 = \lambda$$

$$P(x) = \frac{(1{\cdot}948)^x \cdot e^{-1{\cdot}948}}{x!} \; ; x = 0, 1, 2, \ldots$$

وعن طريق صيغة التراجع يمكن حساب الكتل الاحتمالية المقترنة بعناصر فضاء X وهي :

x :	0	1	2	3
P(x) :	0·1425589	0·2777047	0·2704843	0·1756344
x :	4	5	6	7
P(x) :	0·0855339	0·0333240	0·0108191	0·0003011

عليه فان التكرارات المتوقعة يمكن حسابها من خلالها ضرب $P(x)$ بمجموع التكرارات والموضحة اقيامها في الجدول التالي :

قيم x :	7	6	5	4	3	2	1	0
التكرار المشاهد :	0	0	4	66	90	130	160	50
التكرار المتوقع :	2	5	17	43	88	135	139	71

مثال (٤) : اذا كان $X \sim P0(\lambda)$. جد الدالة المولدة لعزوم الدرجة المعيارية Z . اي $M_Z(t)$ ثم بين ان $\lim_{\lambda \to \infty} M_Z(t) = e^{t^2/2}$

الحل : حيث ان $X \sim P0(\lambda)$ فذلك يعني ان $Z = \frac{X - \lambda}{\sqrt{\lambda}}$

فاذن :

$$M_Z(t) = Ee^{tZ} = Ee^{t\left(\frac{X-\lambda}{\sqrt{\lambda}}\right)}$$

$$= e^{-\sqrt{\lambda}\, t} \cdot M_X\left(\frac{t}{\sqrt{\lambda}} \right)$$

لكن

$$M_X\left(\frac{t}{\sqrt{\lambda}} \right) = e^{\lambda(e^{t/\sqrt{\lambda}} - 1)}$$

عليه فان

$$M_Z(t) = e^{-\sqrt{\lambda}\, t + \lambda(e^{t/\sqrt{\lambda}} - 1)}$$

الان فان المقدار $e^{t/\sqrt{\lambda}}$ وحسب متسلسلة تايلر ماهو الا :

$$e^{t/\sqrt{\lambda}} = \sum_{k=0}^{\infty} \frac{(t/\sqrt{\lambda})^k}{k!} = 1 + \frac{t}{\sqrt{\lambda}} + \frac{t^2}{2\lambda} + 0(\lambda)$$

حيث ان $0(\lambda)$ تعني حدود لاحقة تتضمن λ في مقاماتها بقوى عليا .

وبذلك فان

$$M_Z(t) = e^{-\sqrt{\lambda}\, t + \lambda\left(\frac{t}{\sqrt{\lambda}} + \frac{t^2}{2\lambda} + 0(\lambda) \right)}$$

$$\therefore M_Z(t) = e^{t^2/2 + 0(\lambda)}$$

الان يجعل $\lambda \to \infty$ فذلك يعني ان

$$\lim_{\lambda \to \infty} M_Z(t) = e^{t^2/2}$$

حيث ان الحدود $0(\lambda)$ تتلاشى مقتربة من الصفر عند جعل $\lambda \to \infty$. وسوف نلاحظ لدى دراستنا لموضوع التوزيع الطبيعي المعياري في التوزيعات المستمرة بان الدالة $e^{t^2/2}$ تمثل الدالة المولدة لعزوم هذا التوزيع .

مثال (٥) : افرض ان $X \sim P0(\lambda)$ فاذاعلمت ان $P(X = 1) = 2P(X = 2)$ جد قيمة λ .

الحل : ان

$$P(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \; ; x = 0, 1, 2, \dots$$

عليه فان .

$$P(X = 1) = \lambda e^{-\lambda}, \; P(x = 2) = \frac{\lambda^2 e^{-\lambda}}{2!}$$

وبما ان $P(X=1)=2P(X=2)$ فذلك يعني ان
$$\lambda e^{-\lambda}=\lambda^2 e^{-\lambda}$$

وبحل هذه الصيغة نجد ان $\lambda = 1$.

تمارين عن توزيع پواسون

٥ ـ ٢٤ : افرض ان $X_1 \sim PO(4), X_2 \sim PO(6), X_3 \sim PO(8)$.
وبفرض ان هذه المتغيرات مستقلة تصادفياً . يطلب اجراء مايلي :
أ ـ جد الوسط والتباين الى $Y_1 = X_1 + X_2 , Y_2 = X_1 + X_3$
. $Y_3 = X_1 + X_2 + X_3$
ب ـ جد الدالة المولدة لعزوم $Y = X_1 + X_2 + X_3$
جـ ـ ارسم مخطط دالة الكتلة الاحتمالية الى $Y = X_2 + X_3$.
د ـ جد $P_r(X_1 + X_2 \leq 3), P_r(X_2 + X_3 \geq 2), P_r(X_1 > 3 - X_2)$.

٥ ـ ٢٥ : اذا كان $X \sim PO(\lambda)$. برهن ان :
أ ـ الدالة المولدة للعزوم العاملية هي $M(t) = e^{\lambda(t-1)}$.
ب ـ العزم العاملي ذات المرتبة r هو $\mu_{(r)} = \lambda^r$.

٥ ـ ٢٦ : اذا كان $X = K$ يمثل المنوال الوحيد في توزيع پواسون بالمعلمة λ . برهن ان
$\lambda - 1 \leq K \leq \lambda$

٥ ـ ٢٧ : اذا علمت ان $X \sim P(\lambda)$ وان Y متغير آخر توزيعه الشرطي علماً ان
$X = x$ هو $P(Y=K \mid X=x) = C_k^x P^k (1-P)^{x-k}, K = 0,1,\ldots,x$ و
$0 < P < 1$. برهن ان التوزيع الحدي للمتغير Y هو توزيع پواسون
بالمعلمة (λP) .

٥ ـ ٢٨ : افرض ان $X_1 \sim PO(\lambda_1)$ مستقل عن $X_2 \sim PO(\lambda_2)$. برهن ان
التوزيع الشرطي للمتغير X_1 علماً ان $X_1 + X_2 = n$ هو توزيع ثنائي
الحدين بالمعلمتين $n , P = \lambda_1 / \lambda_1 + \lambda_2$.

٥ ـ ٢٩ : اذا كان $X \sim PO(\lambda)$ ، برهن ان $EX!$ يكون متقارب عندما
$0 < \lambda < 1$ ويكون متباعد عندما $\lambda \geq 1$.

٥ ـ ٧ : توزيع متسلسلة القوى Power series distribution

في هذه الفقرة سوف نستعرض توزيع آخر من النوع المتقطع من شأنه توليد بعض التوزيعات التي سبق وان درسناها في فقرات سابقة .

يقال ان المتغير العشوائي X يتوزع وفق دالة توزيع متسلسلة القوى اذا كانت دالة الكتلة الاحتمالية لهذا المتغير هي

$$P(x) = \frac{a_x . c^x}{f(c)} \quad ; x = 0, 1, 2, \ldots, a_x \geq 0$$

$$= 0 \qquad \text{other wise}$$

حيث a_x دالة غير سالبة بدلالة X وان $f(c)$ دالة موجبة محدودة قابلة للاشتقاق في $c > 0, c$ وان $f(c) = \sum_{x \varepsilon \Omega} a_x c^x$ حيث Ω تمثل مجموعة جزئية معرفة في حقل الاعداد الصحيحة (اي انها مجموعة قابلة للعد) واضح ان

$$\sum_{x \varepsilon \Omega} P(x) = \frac{1}{f(c)} \sum_{x \varepsilon \Omega} a_x . c^x = \frac{1}{f(c)} . f(c) = 1$$

٥ ـ ٧ ـ ١ : الدالة المولدة للعزوم

ان الدالة المولدة لعزوم توزيع متسلسلة القوى هي :

$$M_X(t) = Ee^{tX} = \sum_{x=0}^{\infty} e^{tx} . \frac{a_x . c^x}{f(c)}$$

$$= \frac{1}{f(c)} \sum_{x=0}^{\infty} a_x (ce^t)^x$$

وحيث ان $f(c) = \sum_{x=0}^{\infty} a_x c^x$ فان $\sum_{x=0}^{\infty} a_x (ce^t)^x = f(ce^t)$ فاذن

$$M_X(t) = \frac{f(ce^t)}{f(c)}$$

واضح ان الدالة المولدة التراكمية هي :

$$K_X(t) = \ln M_X(t) = \ln f(ce^t) - \ln f(c)$$

كذلك فان الوسط لهذا التوزيع هو :

$$\mu_x = \sum_{x=0}^{\infty} x \cdot \frac{a_x c^x}{f(c)} = \frac{1}{f(c)} \sum_{x=0}^{\infty} x \cdot a_x \cdot c^x$$

$$= \frac{c}{f(c)} \sum_{x=0}^{\infty} x \cdot a_x \cdot c^{x-1} = \frac{c}{f(c)} \cdot \frac{\partial f(c)}{\partial c}$$

$$= c \cdot \frac{f'(c)}{f(c)}$$

وان التباين في هذا التوزيع هو :

$$\sigma_x^2 = EX^2 - (EX)^2$$

لكن :

$$EX^2 = E[X(X-1) + X] = EX(X-1) + \mu_x$$

$$= \sum_{x=0}^{\infty} x(x-1) \frac{a_x c^x}{f(c)} + \mu_x$$

$$= \frac{c^2}{f(c)} \sum_{x=0}^{\infty} x(x-1) a_x c^{x-2} + \mu_x$$

$$= \frac{c^2}{f(c)} \cdot \frac{\partial^2 f(c)}{\partial c^2} + \mu_x$$

$$= c^2 \frac{f''(c)}{f(c)} + c \frac{f'(c)}{f(c)}$$

فاذن

$$\sigma_x^2 = c^2 \frac{f''(c)}{f(c)} + c \frac{f'(c)}{f(c)} - c^2 \left[\frac{f'(c)}{f(c)} \right]^2$$

٥ ـ ٧ ـ ٢ : حالات خاصة من توزيع متسلسلة القوى .

يمكن استنتاج عدد من التوزيعات المتقطعة التي سبق دراستها في فقرات سابقة كحالات خاصة من توزيع متسلسلة القوى . هذه التوزيعات هي :

١ـ توزيع ثنائي الحدين :

اذا تم اختيار $c = \frac{P}{q}$ حيث $f(c) = (1+c)^n, q = 1 - P, 0 < P < 1$

حيث n عدد موجب صحيح ، وان $\Omega = \{x : x = 0, 1, 2, \ldots, n\}$

عندئذ :

$$f(c) = \sum_{x=0}^{n} a_x c^x = (1+c)^n$$

وحسب نظرية ثنائي الحدين فان

$$(1+c)^n = \sum_{x=0}^{n} c_x^n . c^x$$

فاذن

$$f(c) = \sum_{x=0}^{n} a_x c^x = \sum_{x=0}^{n} c_x^n . c^x$$

وهذا يعني ان $a_x = c_x^n$ عليه فان :

$$P(x) = \frac{a_x c^x}{f(c)} = \frac{c_x^n (P/q)^x}{\left(1 + \frac{P}{q}\right)^n}$$

$$= c_x^n . P^x . q^{-x} . q^n = c_x^n P^x . q^{n-x}, x = 0, 1, 2, \ldots, n$$

لاحظ ان

$$M_X(t) = \frac{f(ce^t)}{f(c)} = \frac{(1+ce^t)^n}{(1+c)^n}$$

$$= \frac{\left(1 + \frac{P}{q} e^t\right)^n}{\left(1 + \frac{P}{q}\right)^n} = q^n\left(1 + \frac{P}{q} e^t\right)^n$$

$$\therefore M_X(t) = (q + Pe^t)^n$$

ومنها يمكن حساب عزوم التوزيع .

٢ـ توزيع ثنائي الحدين السالب .

بفرض ان $c = \frac{P}{1+P}$ حيث $0 < P < 1$ وان $f(c) = (1 - c)^{-r}$ حيث r عدد موجب، $\Omega = \{x : x = 0, 1, 2, \ldots\}$ عندئذٍ فان

$$f(c) = \sum_{x=0}^{\infty} a_x c^x = (1 - c)^{-r}$$

وحسب مفكوك ثنائي الحدين السالب فان

$$(1 - c)^{-r} = \sum_{x=0}^{\infty} (-1)^x . c_x^{-r} . c^x$$

فاذن

$$f(c) = \sum_{x=0}^{\infty} a_x c^x = \sum_{x=0}^{\infty} (-1)^x c_x^{-r} c^x$$

وهذا يعني ان $a_x = (-1)^x . c_x^{-r}$ عليه فان .

$$P(x) = \frac{a_x c^x}{f(c)} = \frac{(-1)^x . C_x^{-r} (P/1+p)^x}{\left(1 - \frac{P}{1+P}\right)^{-r}}$$

$$= (-1)^x . c_x^{-r}\left(\frac{P}{1+P}\right)^x . \left(\frac{1}{1+P}\right)^r$$

$$\therefore P(x) = c_x^{-r}\left(\frac{1}{1+P}\right)^r.\left(-\frac{P}{1+P}\right)^x ; x = 0, 1, 2, \ldots$$

فاذن

$$X \sim Nb\left(r, \frac{1}{1+P}\right)$$

ويطلب من القاريء ايجاد الدالة المولدة لعزوم هذا التوزيع .

٣ـ توزيع پواسون

بفرض ان $f(c) = e^c$ وان $\Omega = \{x : x = 0, 1, 2, \ldots\}$ عندئذٍ فان

$$f(c) = \sum_{x=0}^{\infty} a_x c^x = e^c$$

وحسب متسلسلة تايلر فان

$$e^c = \sum_{x=0}^{\infty} \frac{c^x}{x!}$$

لذا فان

$$f(c) = \sum_{x=0}^{\infty} a_x c^x = \sum_{x=0}^{\infty} \frac{1}{x!} . c^x$$

يعني ان $a_x = \frac{1}{x!}$. فاذن

$$P(x) = \frac{a_x c^x}{f(c)} = \frac{1}{x!} . \frac{c^x}{e^c}$$

$$= \frac{c^x e^{-c}}{x!} ; x = 0, 1, 2, \ldots$$

فاذن $X \sim P0(c)$.

٤ ـ توزيع المتسلسلة اللوغارتمية
Logarithmic series distribution

بفرض ان $f(c) = -\ln(1-c)$ وان $\Omega = \{x : x = 1, 2, \ldots\}$ عندئذٍ

$$f(c) = \sum_{x=1}^{\infty} a_x c^x = -\ln(1-c)$$

وحسب متسلسلة تايلر فان

$$-\ln(1-c) = \sum_{x=1}^{\infty} \frac{c^x}{x}, 0 < c < 1$$

فاذن

$$f(c) = \sum_{x=1}^{\infty} a_x c^x = \sum_{x=1}^{\infty} \frac{1}{x} . c^x$$

وهذا يعني ان $a_x = \frac{1}{x}$. لذا فان :

$$P(x) = \frac{a_x c^x}{f(c)} = \frac{c^x}{-x\ln(1-c)}$$

$$= \frac{C^x}{\ln(1-c)^{-x}} ; x = 1, 2, \ldots, 0 < C < 1$$

واذا تم اختيار $C = P$ حيث P يمثل احتمال نجاح المحاولة في تجربة مافان

$$P(x; P) = \frac{P^x}{\ln q^{-x}} ; x = 1, 2, \ldots, 0 < P < 1, q = 1 - P$$

تمارين عن توزيع متسلسلة القوى

٥ ـ ٣٠ : افرض ان X متغير عشوائي يتوزع وفق دالة توزيع متسلسلة القوى وان $x = 0, 1, 2, ..., f(c) = (1 - c)^{-1}, 0 < P < 1$ حيث $C = P/(1 + P)$ ماهو التوزيع الاحتمالي للمتغير X وفق هذه المعطيات؟ جد الوسط والتباين لهذا المتغير.

٥ ـ ٣١ : برهن ان الدالة المولدة للعزوم العاملية لتوزيع متسلسلة القوى هي :

$$M(t) = \frac{f[c(1 + t)]}{f(c)}$$

٥ ـ ٣٢ : برهن ان العزم العاملي ذا المرتبة K في توزيع متسلسلة القوى معطى بالصيغة

$$\mu_{(k)} = C^k f^{(k)}(c)/f(c)$$

٥ ـ ٣٣ : برهن ان صيغة التراجع في توزيع متسلسلة القوى هي :

$$P(x + 1) = C.\frac{a_{x+1}}{a_x} \cdot P(x).$$

٥ ـ ٨ : التوزيع متعدد الحدود
The Multinomial distribution

لاحظنا في الفقرات السابقة من هذا الفصل ان دراستنا للتوزيعات المتقطعة كانت منصبة على حالة وجود متغير عشوائي واحد uni - variate يتوزع وفق دالة كتلة احتمالية كدالة توزيع ثنائي الحدين ودالة توزيع بواسون وغيرها . في هذه الفقرة سوف نستعرض حالة وجود عدة متغيرات تتوزع مجتمعة وفق دالة كتلة احتمالية مشتركة . ان احد اهم التوزيعات المتعددة المتغيرات ومن النوع المتقطع هو توزيع متعدد الحدود الذي يمكن اعتباره حالة اكثر عمومية لتوزيع ثنائي الحدين . علماً ان هنالك توزيعات اخرى متعددة المتغيرات من النوع المتقطع مثل توزيع بواسون متعدد المتغيرات وتوزيع ثنائي الحدين متعدد المتغيرات وغيرها . وفيما يلي تعريف لهذا التوزيع : يقال ان المتغيرات $X_1, X_2, \ldots, X_k$ تتوزع مجتمعة وفق دالة توزيع متعدد الحدود اذا كانت دالة الكتلة الاحتمالية المشتركة لهذه المتغيرات تأخذ الشكل التالي :

$$P(x_1, x_2, \ldots, x_k) = \frac{n!}{x_1!\, x_2! \ldots x_k!} \cdot P_1^{x_1} \cdot P_2^{x_2} \ldots\ldots P_k^{x_k}$$

حيث ان $x_i = 0, 1, 2, \ldots, n \,;\, i = 1, 2, \ldots, k$ وان $n, P_1, P_2, \ldots, P_k$ تمثل معالم التوزيع بحيث ان $0<P_i<1$ وان $\sum_{i=1}^{k} X_i = n \,,\; \sum_{i=1}^{k} P_i = 1$ وللسهولة في الكتابة وبفرض ان $x' = [\,x_1\; x_2 \ldots x_k\,]$ يمثل متجه vector صفي عناصره تمثل المتغيرات $X_1, X_2, \ldots, X_k$ فان

$$P(x) = \frac{n!}{\prod_{i=1}^{k} x_i!} \cdot \prod_{i=1}^{k} P_i^{x_i}$$

ان الدالة $P(x)$ في الحقيقة لا ماهي الا الحد العام لمفكوك متعدد الحدود للصيغة $(P_1 + P_2 + \dots P_k)^n$ وعندما $K = 2$ فان :

$$P(x_1, x_2) = \frac{n!}{x_1!\, x_2!} P_1^{x_1} . P_2^{x_2}$$

لكن $x_1 + x_2 = n$ فاذن $x_1 = n - x_2$ وان $P_1 + P_2 = 1$ فاذن $P_1 = 1 - P_2$ وهذا يعني ان

$$P(x_2) = \frac{n!}{x_2!\,(n - x_2)!} P_2^{x_2} (1 - P_2)^{n - x_2};\ x_2 = 0, 1, 2, \dots, n.$$

$$= C_{x_2}^n P_2^{x_2} (1 - P_2)^{n - x_2}$$

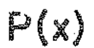
او ان

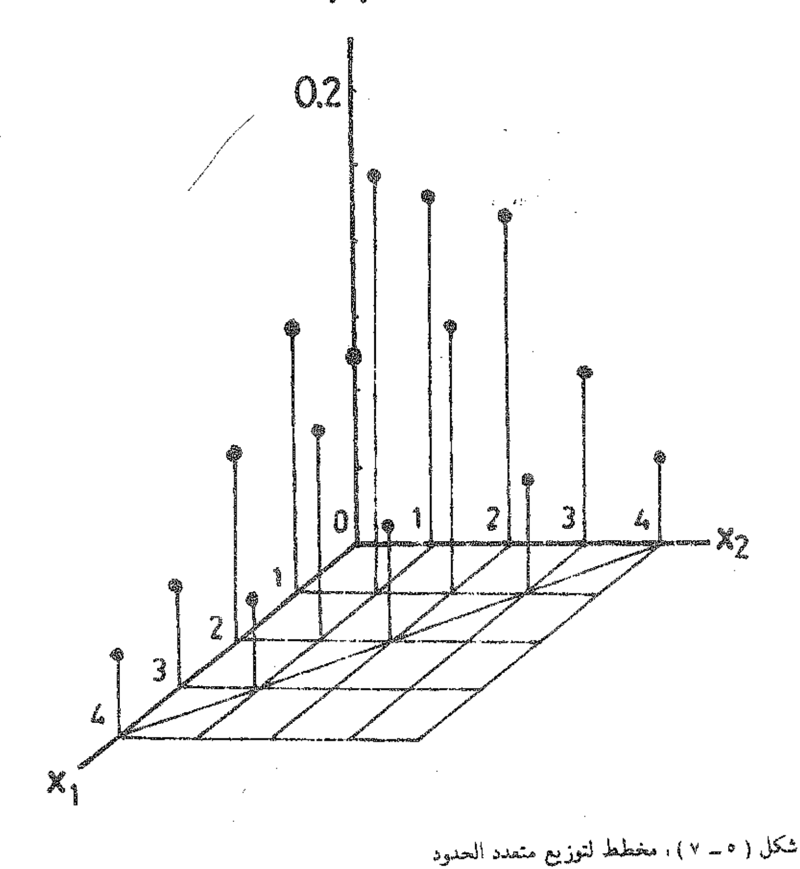
$$P(x_1) = C_{x_1}^n P_1^{x_1} (1 - P_1)^{n - x_1};\ x_1 = 0, 1, 2, \dots, n$$

من ذلك يتضح انه في حالة وجود متغيرين فقط فان دالة هذا التوزيع تعبر عن دالة توزيع ثنائي الحدين للمتغير X_1 او المتغير X_2. فمثلاً اذا كان X_1 يمثل عدد حالات النجاح في تجربة معينة فيها احتمال نجاح المحاولة الواحدة هو P_1 وان $X_1 \sim b(n, P_1)$ فان عدد حالات الفشل X_2 في هذه التجربة هو ايضاً متغير عشوائي يتوزع وفق دالة توزيع ثنائي الحدين بالمعلمتين n . $P_2 = 1 - P_1$ اي انه يمكن النظر لهذه التجربة بمنظارين اما عدد حالات نجاح التجربة او عدد حالات فشلها فاذا علم عدد حالات النجاح يمكن التوصل لعدد حالات الفشل والعكس صحيح طالما ان مجموعهما مساوٍ لعدد المحاولات n وان احتمال النجاح مضاف اليه احتمال الفشل هو العدد (1) . والشكل (٥ ــ ٧) يوضح مخطط دالة توزيع متعدد الحدود بالمعالم 4, 0·2, 0·3, 0·5 .

ويمكن بيان ان مجموع الكتل الاحتمالية المشتركة المقترنة بعناصر المتجه x مساوٍ للواحد دلالة على ان دالة توزيع متعدد الحدود هي دالة كتلة احتمالية مشتركة كما هو مبين بالآتي :

$$\sum_x P(x) = \sum_x \frac{n!}{\prod_{i=1}^{k} x_i!} . \prod_{i=1}^{k} P_i^{x_i}$$





شكل (٥ ـ ٧) ، مخطط لتوزيع متعدد الحدود

$$P(x) = \frac{4!}{x_1!x_2!(4-x_1-x_2)!}(0.2)^{x_1}.(0.3)^{x_2}.\ (0.5)^{4-x_1-x_2}$$

لكن وطبقاً لنظرية متعدد الحدود من درجة n فان :

$$(P_1 + P_2 + \dots + P_k)^n = \sum_x \frac{n!}{\prod_{i=1}^{k} x_i!} . \prod_{i=1}^{k} P_i^{x_i}$$

لذا فان :

$$\sum_x P(x) = (P_1 + P_2 + \dots + P_k)^n = 1^n = 1$$

٥ ــ ٨ ــ ١ : الدالة المولده لعزوم توزيع متعدد الحدود .

ان الدالة المولدة لعزوم توزيع متعدد الحدود من شأنها توليد العزوم الحدية (اي عزوم كل متغير بشكل منفرد) وكذلك العزوم المشتركة ما بين اي متغيرين او اكثر من متغيرات المتجه x . هذه الدالة هي $\left[\sum_{i=1}^{k} P_i e^{t_i}\right]^n$

البرهان :

$$M_x(t) = Ee^{t'X}$$

حيث $t' = [\, t_1 t_2 \dots t_k \,]$

$$= \sum_x e^{\sum_{i=1}^{k} t_i x_i} \cdot \frac{n!}{\prod_{i=1}^{k} x_i!} \cdot \prod_{i=1}^{k} P_i^{x_i} = \sum_x \frac{n!}{\prod_{i=1}^{k} x_i!} \cdot \prod_{i=1}^{k} (P_i e^{t_i})^{x_i}$$

وطبقاً لنظرية متعدد الحدود من درجة n ولأية اعداد حقيقة مثل $a_1, a_2, \dots, a_k$ فان :

$$(a_1 + a_2 + \dots + a_k)^n = \sum_x \frac{n!}{\prod_{i=1}^{k} x_i!} \cdot \prod_{i=1}^{k} a_i^{x_i}$$

وبوضع $a_i = P_i e^{t_i}$ نحصل على :

$$(P_1 e^{t_1} + P_2 e^{t_2} + \dots + P_k e^{t_k})^n = \sum_x \frac{n!}{\prod_{i=1}^{k} x_i!} \prod_{i=1}^{k} (P_i e^{t_i})^{x_i}$$

فاذن

$$M_X(t) = \left[\sum_{i=1}^{k} P_i e^{t_i}\right]^n$$

ويتضح من هذه الدالة مايلي :

$1 - M_X(t = 0) = 1$

$2 - M_X(t = 0 \text{ accept } t_i \neq 0) = M_{X_i}(t_i)$

$3 - M_X(t = 0 \text{ accept } t_i, t_j \neq 0) = M_{X_i X_j}(t_i, t_j)$

$$4 - K_X(t) = n \ln \sum_{i=1}^{k} P_i e^{t_i}$$

٥ ـ ٨ ـ ٢ : الوسط والتباين والتباين المشترك

لاغراض السهولة في التوصل لصيغ هذه المقاييس فاننا سوف نستخدم الدالة المولدة التراكمية $K_X(t)$ وكما يلي :

ان

$$\mu_{x_i} = \frac{\partial K_X(t)}{\partial t_i}\Bigg]_{t=0}$$

لكن

$$\frac{\partial K_X(t)}{\partial t_i} = \frac{nP_i e^{t_i}}{\sum_{i=1}^{k} P_i e^{t_i}}$$

وبجعل المتجه t مساوٍ للمتجه الصفري (اي ان $t_1 = t_2 = \ldots = t_k = 0$) نحصل على

$$\mu_{x_i} = EX_i = nP_i, \; i = 1, 2, \ldots, k$$

كذلك فان

$$\sigma^2_{x_i} = \frac{\partial^2 K_X(t)}{\partial t_i^2}\Bigg]_{t=0}$$

لكن

$$\frac{\partial^2 K_X(t)}{\partial t_i^2} = nP_i \left[\frac{\left(\sum_{i=1}^{k} P_i e^{t_i}\right)(e^{t_i}) - (e^{t_i})(P_i e^{t_i})}{\left(\sum_{i=1}^{k} P_i e^{t_i}\right)^2} \right.$$

وبجعل المتجه t مساوٍ للمتجه الصفرى نحصل على :

$$\sigma^2_{x_i} = nP_i(1 - P_i), \; i = 1, 2, \ldots, k$$

كما وان

$$\sigma_{ij} = \frac{\partial^2 K_X(t)}{\partial t_i \cdot \partial t_j}\Bigg]_{t=0}$$

لكن

$$\frac{\partial^2 K_X(t)}{\partial t_i \partial t_j} = nP_i e^{t_i}\left[-\left(\sum_{i=1}^{k} P_i e^{t_i}\right)^{-2} . P_j e^{t_j}\right]$$

وبوضع $t = 0$ نحصل على

$$\sigma_{ij} = -nP_iP_j \ , i , j = 1 , 2 , \ldots , n$$

٥ ـ ٨ ـ ٣ : مصفوفة التباين والتباين المشترك ومصفوفة الارتباطات .

على ضوء صيغ التباين والتباين المشترك التي حصلنا عليها في الفقرة (٥ ـ ٨ ـ ٢) يمكن تكوين مصفوفة التباين والتباين المشترك للمتجه X وكذلك مصفوفة الارتباطات بين اي متغيرين من متغيرات هذا المتجه وعلى هذا النحو الآتي :

بفرض ان $\sigma_i^2 = \sigma_{x_i}^2 = nP_i(1 - P_i)$ وان $Cov(X_i, X_j) = \sigma_{ij}$ فان مصفوفة التباين والتباين المشترك للمتجه X هي :

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2k} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \cdots & \sigma_{3k} \\ \vdots & \vdots & \vdots & & \vdots \\ \sigma_{k1} & \sigma_{k2} & \sigma_{k3} & \cdots & \sigma_k^2 \end{bmatrix} , \begin{matrix} \sigma_i^2 = nP_i(1 - P_i) \\ \sigma_{ij} = -nP_iP_j \end{matrix}$$

علماً ان المصفوفة V مصفوفة متماثلة $(\sigma_{ij} = \sigma_{ji})$ ذات مرتبة KXK

كذلك فان :

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i . \sigma_j} = \frac{-nP_iP_j}{\sqrt{nP_iq_i} . \sqrt{nP_jq_j}}$$

$$= -\sqrt{\frac{P_iP_j}{q_iq_j}}$$

وعندئذٍ فان مصفوفة الارتباطات R ستكون .

$$R = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2k} \\ \rho_{31} & \rho_{32} & 1 & \cdots & \rho_{3k} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & \rho_{k3} & \cdots & 1 \end{bmatrix}$$

كذلك فان R مصفوفة متماثلة $(\rho_{ij} = \rho_{ji})$ ذات مرتبة $K \times K$. ان رتبه rank المصفوفة V وكذلك المصفوفة R هي $(K - 1)$ بسبب ان $|V| = |R| = 0$ (الرمز | | يعني محدد المصفوفة) . وهذا يعني ان هاتين المصفوفتين شاذتان Singular matrices . وعلى هذا الاساس يقال ان توزيع متعدد الحدود هو توزيع شاذ Singular distribution وان السبب في ذلك يعود للشرط المفروض على هذا التوزيع وهو ان $\sum_{i=1}^{k} P_i = 1 , \sum_{i=1}^{k} X_i = n$ لذلك ومن الناحية العملية فانه غالباً ما يتم حذف احد المتغيرات الاقل اهمية من بين متغيرات المتجه X والتعامل مع $(k - 1)$ متغير فقط حيث ان هذا الاجراء يؤدي الى التخلص من حالة الشذوذ في هذا التوزيع .

٥ ـ ٨ ـ ٤ : مثال : اذا كانت :

$$P(x_1, x_2, x_3) = \frac{10!}{x_1!\, x_2!\, x_3!} (0{\cdot}1)^{x_1} . (0{\cdot}3)^{x_2} . (0{\cdot}6)^{x_3}$$

تمثل دالة توزيع متعدد الحدود بالمتغيرات X_1, X_2, X_3 عندئذٍ :

١ ـ فان متوسطات هذه المتغيرات هي :

$$EX_1 = (10)(0{\cdot}1) = 1, EX_2 = (10)(0{\cdot}3) = 3, EX_3 = (10)(0{\cdot}6) = 6$$

٢ ـ ان تباينات هذه المتغيرات هي :

$$\sigma_1^2 = 10(0{\cdot}1)(0{\cdot}9) = 0{\cdot}9 \quad , \quad \sigma_2^2 = 10(0{\cdot}3)(0{\cdot}7) = 2{\cdot}1 ,$$
$$\sigma_3^2 = 10(0{\cdot}6)(0{\cdot}4) = 2{\cdot}4$$

٣ ـ ان التباينات المشتركة ما بين اي متغيرين هي :

$$\sigma_{12} = -(10)(0{\cdot}1)(0{\cdot}3) = -0{\cdot}3,\ \sigma_{13} = -(10)(0{\cdot}1)(0{\cdot}6) = -0{\cdot}6,$$
$$\sigma_{23} = -(10)(0{\cdot}3)(0{\cdot}6) = -1{\cdot}8$$

٤ ـ ان مصفوف التباين والتباين المشترك هي :

$$V = \begin{bmatrix} 0{\cdot}9 & -0{\cdot}3 & -0{\cdot}6 \\ -0{\cdot}3 & 2{\cdot}1 & -1{\cdot}8 \\ -0{\cdot}6 & -1{\cdot}8 & 2{\cdot}4 \end{bmatrix}$$

لاحظ ان

$$|V| = \begin{vmatrix} 0{\cdot}9 & -0{\cdot}3 & -0{\cdot}6 \\ -0{\cdot}3 & 2{\cdot}1 & -1{\cdot}8 \\ -0{\cdot}6 & -1{\cdot}8 & 2{\cdot}4 \end{vmatrix} = 0$$

٥ ـ ان معاملات الارتباط بين اي متغيرين من هذه المتغيرات هي :

$$\rho_{12} = -0{\cdot}2182,\ \rho_{13} = -0{\cdot}4082,\ \rho_{23} = -0{\cdot}8018.$$

فاذن :

$$R = \begin{bmatrix} 1 & -0{\cdot}2182 & -0{\cdot}4082 \\ -0{\cdot}2182 & 1 & -0{\cdot}8018 \\ -0{\cdot}4082 & -0{\cdot}3018 & 1 \end{bmatrix},\ |R| = 0$$

٥ ــ ٣٤ : اذا علمت ان المتغيرات العشوائية $X_1, X_2, \dots, X_k$ تتوزع مجتمعة وفق دالة توزيع متعدد الحدود بالمعالم $n, P_1, P_2, \dots, P_k$ برهن ان $X_i \sim b(n, P_i)$.

٥ ــ ٣٥ : اذا كانت X_1, X_2, X_3 متغيرات عشوائية تتوزع وفق توزيع متعددالحدود بالمعالم n, P_1, P_2, P_3 . جد التوزيع المشترك الشرطي الى X_1, X_2 علماً ان $X_3 = x_3$.

٥ ــ ٣٦ : افرض ان $X_1, X_2, \dots, X_k$ متغيرات عشوائية مستقلة بحيث أن $X_i \sim Po(\lambda_i)$. برهن ان التوزيع الشرطي لحادثة تقاطع (حاصل ضرب) هذه المتغيرات علماً ان $\sum_{i=1}^{k} X_i = n$ هو توزيع متعدد الحدود بالمعالم n ، $P_i = \frac{\lambda_i}{\sum_{i=1}^{k} \lambda_i}$. ثم بين اذا كانت $\lambda_i = \lambda$ ، $i = 1, 2, \dots, k$. فان هذا التوزيع الشرطي هو توزيع متعدد الحدود بالمعالم n ، $P_i = \frac{1}{k}$.

٥ ــ ٣٧ : اذا علمت ان X_1, X_2, X_3, X_4 متغيرات عشوائية تتوزع وفق دالة توزيع متعدد الحدود بالمعالم $20, 0.1, 0.2, 0.3, 0.4$ على التوالي . جد مايلي :

أ ــ دالة التوزيع المشترك لهذه المتغيرات .

ب ــ الوسط والتباين لكل متغير منها وكذلك التباين المشترك .

جـ ــ مصفوفة التباين والتباين المشترك ومصفوفة الارتباطات ثم بين ان هاتان المصفوفتان شاذتان .

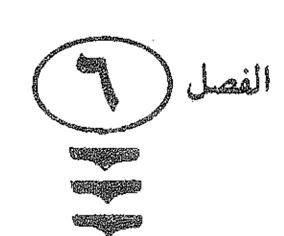
د ــ معامل الارتباط الجزئي بين X_1, X_2 بعد استبعاد اثر المتغير X_3 ، معامل الارتباط الجزئي بين X_1, X_2 بعد استبعاد اثر X_3, X_4

هـ ــ معامل الارتباط المتعدد بين X_1 مع X_2, X_3, X_4 .

و ــ التوزيع الحدي لكل متغير من هذه المتغيرات .

ز ــ التوزيع المشترك للمتغرين X_1, X_2 . ماهو هذا التوزيع ؟





التوزيعات المستمرة النظرية

الفصل السادس

التوزيعات المستمرة النظرية

استعرضنا في الفصل الخامس اهم التوزيعات المتقطعة النظرية ذات الاهمية التطبيقية في النظرية الاحصائية . في هذا الفصل سوف نركز الاهتمام على دراسة أهم التوزيعات المستمرة النظرية وذلك من خلال اعطاء تعريف متكامل لكل توزيع مع عرض لاهم خصائصه .

٦ ـ ١ : التوزيع المنتظم المستمر
Continuous Uniform distribution

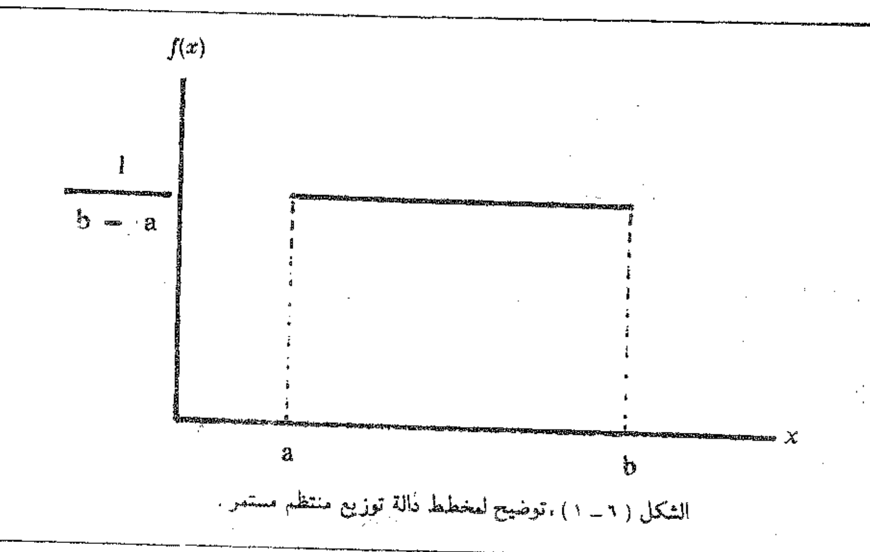
ويسمى في بعض الاحيان بالتوزيع المستطيل Rectangular distribution نظراً لان مخطط دالة هذا التوزيع يأخذ شكل المستطيل . ان اهم استخدامات هذا التوزيع هي تكوين مايسمى بـ « جداول الاعداد العشوائية » التي تستخدم في اختيار عينة عشوائية من مجتمع احصائي .

يقال ان المتغير العشوائي X يتوزع وفق دالة التوزيع المنتظم المستمر اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

$$f(x; a, b) = \frac{1}{b - a} \; ; a \leq x \leq b$$

$$= 0 \qquad \text{other wise}$$

حيث a, b هما معلمتا هذا التوزيع وان b > a عددان حقيقيان . والشكل (٦ ـ ١) يوضح مخطط دالة هذا التوزيع .



الشكل (٦ ـ ١) ، توضيح لمخطط دالة توزيع منتظم مستمر .

ويمكن بيان ان المساحة تحت مخطط الدالة $f(x)$ المحدودة بالمستقيمين $x = b , x = a$ مساوية للواحد دلالة على كون $f(x)$ دالة كثافة احتمالية وكالآتي :

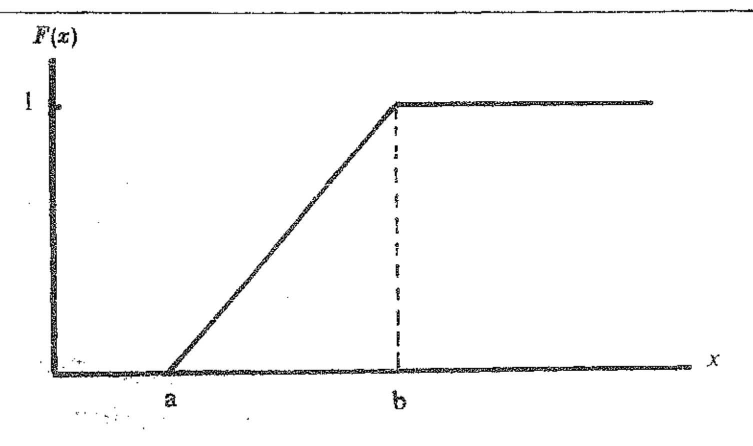
$$\int_a^b f(x)\,dx = \frac{1}{b-a}\int_a^b dx = 1$$

٦ ـ ١ ـ ١ : الدالة التوزيعية Distribution function

ان الدالة التوزيعية في التوزيع المنتظم المستمر هي :

$$F(x) = \int_a^x f(u)\,du = \frac{1}{b-a}\int_a^x du = \frac{x-a}{b-a} \; ; a \le x \le b$$

واضح ان $F(b) = 1 , F(a) = 0$. والشكل (٦ ـ ٢) يوضح مخطط هذه الدالة .



شكل (٦ ـ ٢) ، مخطط الدالة $F(x)$ لتوزيع منتظم مستمر

٦ ـ ٢ ـ ١ : الوسط والتباين Mean and Variance

ان الوسط في التوزيع المنتظم المستمر هو $\frac{a+b}{2}$ والتباين هو $\frac{(b-a)^2}{12}$

البرهان :

$$\mu_x = \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\cdot\frac{b^2-a^2}{2}$$

$$= \frac{a+b}{2}$$

كذلك فان

$$\sigma_x^2 = EX^2 - \mu_x^2$$

وبما ان

$$EX^2 = \frac{1}{b-a}\int_a^b x^2dx = \frac{1}{b-a}\cdot\frac{b^3-a^3}{3}$$

$$= \frac{a^2+ab+b^2}{3}$$

فاذن

$$\sigma_x^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2} \right)^2 = \frac{(b-a)^2}{12}$$

٦ ـ ١ ـ ٢ : الدالة المولدة للعزوم

Moment generating function

يمتلك التوزيع المنتظم المستمر دالة مولدة للعزوم . هذه الدالة هي :

$$M_X(t) = Ee^{tX} = \frac{1}{b-a} \int_a^b e^{tx} dx$$

$$= \frac{e^{tb} - e^{ta}}{t(b-a)} \quad , t > 0$$

ويتضح من هذه الدالة أن $M_X(0) = \frac{0}{0}$ وهو شكل غير محدد ومن المعروف ان $M_X(0) = 1$. لكن وباستخدام « قاعدة لوبتيل Lopital's Rule » يمكن اثبات ان $M_X(0) = 1$ وكالآتي :

$$M_X(t) = \frac{e^{tb} - e^{ta}}{t(b-a)} = \frac{g_1(t)}{g_2(t)}$$

فاذن $g_1'(t) = be^{tb} - ae^{ta}$, $g_2'(t) = b - a$ عليه فان

$$\lim_{t \to 0} M_X(t) = \lim_{t \to 0} \frac{g_1'(t)}{g_2'(t)} = \frac{b-a}{b-a} = 1$$

وحيث ان مسألة التعامل مع هذه الدالة لتوليد عزوم التوزيع تبدو معقدة لذلك سوف نستعيض عن هذه الدالة من خلال اشتقاق صيغة للعزم ذي المرتبة r حول نقطة الاصل وكما يلي :

$$EX^r = \frac{1}{b-a}\int_a^b x^r dx = \frac{b^{r+1}-a^{r+1}}{(r+1)(b-a)} = M_X^{(r)}(0) ; r = 1, 2, \ldots$$

٦ ـ ١ ـ ٤ : العزوم المركزية

ان العزم ذو المرتبة r في التوزيع المنتظم المستمر هو :

$$E(X-\mu_x)^r = \frac{1}{b-a}\int_a^b (x-\mu_x)^r dx$$

$$= \frac{(b-\mu_x)^{r+1}-(a-\mu_x)^{r+1}}{(r+1)(b-a)}$$

وبما ان

$$b-\mu_x = b - \frac{a+b}{2} = \frac{b-a}{2}$$

وان

$$a-\mu_x = a - \frac{a+b}{2} = \frac{a-b}{2}$$

وبالتعويض واخراج عامل مشترك نحصل على .

$$E(X-\mu_x)^r = \frac{(b-a)^{r+1}.[1-(-1)^{r+1}]}{2^{r+1}.(r+1)(b-a)}$$

ويتضح من هذه الصيغة ان

$$E(X-\mu_x)^r = \frac{(b-a)^r}{2^r(r+1)} \quad ; \quad r \text{ عدد زوجي}$$

$$= 0 \quad ; \quad r \text{ عدد فردي}$$

٦ ـ ١ ـ ٥ : خاصة البتر في التوزيع المنتظم المستمر .

بفرض ان X يتوزع وفق دالة توزيع منتظم مستمر على الفترة (a,b). عندئذٍ فان اي بتر في التوزيع يولد توزيع آخر هو توزيع منتظم مستمر .

البرهان :

افرض اننا نرغب بعمل بتر في التوزيع بحيث ان الدالة $f(x)$ تكون معرفة على الفترة (c,d) ، $a < c, d < b, c < d$ عندئذٍ فان

$$f(x \mid c < x < d) = \frac{f(x)}{F(d) - F(c)}$$

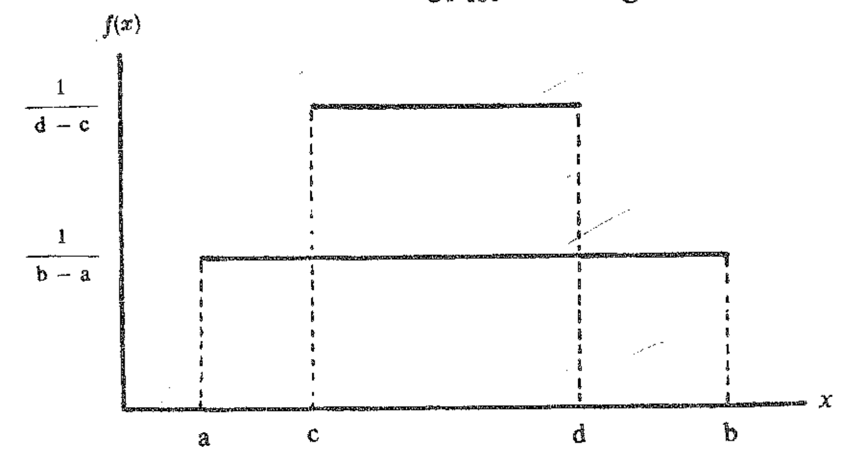
لكن

$$F(d) = \frac{d-a}{b-a}, F(c) = \frac{c-a}{b-a}$$

فاذن

$$f(x \mid c < x < d) = \frac{\frac{1}{b-a}}{\frac{d-a}{b-a} - \frac{c-a}{b-a}} = \frac{1}{d-c}, \quad c \leq x \leq d$$

والدالة الأخيرة ماهي الا دالة توزيع منتظم مستمر معرفة على الفترة (c,d) والشكل (٦ ـ ٣) يوضح مخطط التوزيعين :



الشكل (٦ ـ ٣) : توضيح لمخطط توزيعين منتظمين .

٦ ـ ١ ـ ٦ : أمثلة :

مثال (١) : اذا كان X يتوزع كتوزيع منتظم مستمر على الفترة (6 و 2) عندئذ

$$1 - \ f(x) = \frac{1}{4} \ ; 2 \le x \le 6$$

$$2 - \ F(x) = \frac{x-2}{4} \ ; 2 \le x \le 6$$

$$3 - \ \mu_x = 4 \, , \sigma_x^2 = \frac{4}{3}$$

$$4 - \ EX^r = \frac{6^{r+1} - 2^{r+1}}{4(r+1)} \, , r = 1, 2, \ldots$$

$$5 - \ E(X - \mu_x)^r = \frac{2^r}{(r+1)} \, , r = 2, 4, 6, \ldots$$

$$6 - \ P_r(X < 5) = F(5) = 0.75$$

مثال (٢) : اذا علمت ان X متغير عشوائي يتوزع كتوزيع منتظم مستمر على الفترة $(-a, a)$، $a > 0$ جد قيمة a بحيث ان $P_r(X > 1) = \frac{1}{3}$

الحل :

واضح هنا ان

$$f(x) = \frac{1}{2a} \, , -a < x < a$$

عليه فان

$$P_r(X > 1) = \int_1^a \frac{1}{2a} \, dx = \frac{1}{3}$$

او ان

$$\frac{1}{2a}(a - 1) = \frac{1}{3}$$

وبحل الصيغة الاخيرة نجد ان $a = 3$.

مثال (٣) : اذا علمت ان X يتوزع كتوزيع منتظم مستمر على الفترة (a , b) وافرض ان $Q_j , j = 1, 2, 3$ يمثل الربيع j . وان $D_j = j = 1, 2, \ldots, 9$ يمثل العشير j . برهن ان

$$Q_j = \frac{(4-j)a + jb}{4} , D_j = \frac{(10-j)a + jb}{10}$$

البرهان : حيث ان X يتوزع كتوزيع منتظم مستمر على الفترة (a , b). فذلك يعني ان :

$$F(x) = \frac{x-a}{b-a}$$

وبذلك فان الربيع j · يمثل قيمة X التي تجعل $F(x) = \frac{j}{4}$. عليه فان

$$F(x) = \frac{x-a}{b-a} = \frac{j}{4}$$

وبحل هذه الصيغة نسبة الى X ومن ثم التعويض عن X بـ Q_j نحصل على :

$$Q_j = \frac{(4-j)a + jb}{4} , j = 1, 2, 3$$

ووفق نفس الاجراء نترك برهنة الشق الثاني من المثال للقاريء .

تمارين عن التوزيع المنتظم المستمر

٦ ـ ١ : اذا علمت ان X يتوزع كتوزيع منتظم مستمر على الفترة (10 و 4) . جد ما يلي :

أ ـ دالة الكثافة الاحتمالية للمتغير X ثم ارسم مخطط هذه الدالة .

ب ـ الدالة التوزيعية مع رسم مخططها .

جـ ـ الوسط والتباين في هذا التوزيع .

د ـ العزم الثالث والرابع حول نقطة الاصل والعزم المركزي الرابع .

٦ ـ ٢ : اذا كان X يتوزع كتوزيع منتظم مستمر على الفترة (a , b) . برهن أن :

أ ـ الوسيط لهذا التوزيع مساو لمتوسطه .

ب ـ الانحراف المتوسط لهذا التوزيع هو $\frac{1}{4}(b - a)$.

٦ ـ ٣ : ليكن X متغيراً عشوائياً يتوزع كتوزيع منتظم مستمر على الفترة (a , a −) $a > 0$ برهن أن :

أ ـ الدالة المولدة لعزوم X حول نقطة الاصل هي $M_X(t) = (\sinh at)/at$.

ب ـ الدالة المميزة لهذا التوزيع هي $\phi(t) = e^{ita} . \sin at / at$

جـ ـ ان العزم المركزي ذا المرتبة r (r عدد زوجي) هو $ar(r + 1)^{-1}$

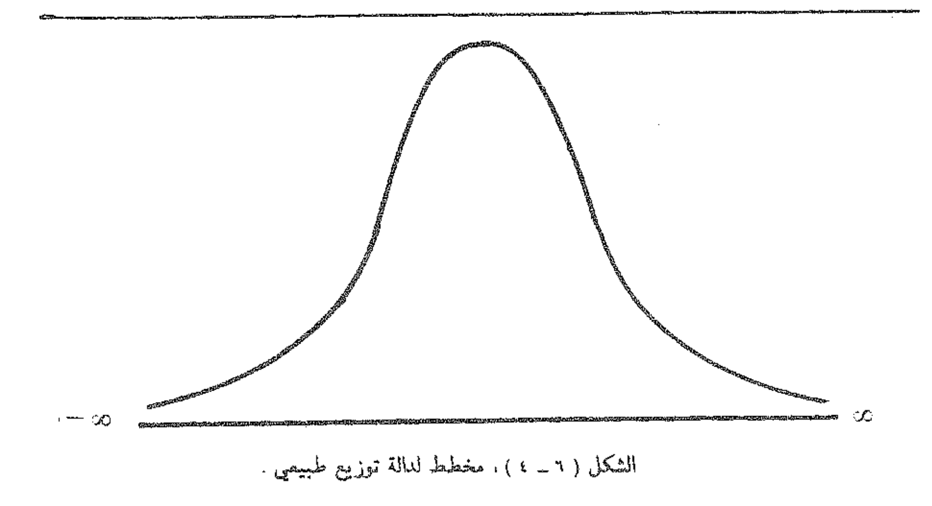
٦ ـ ٤ : اشتق صيغة لمعامل الاختلاف في التوزيع المنتظم المستمر .

٦ ـ ٢ : التوزيع الطبيعي Normal distribution

يعتبر التوزيع الطبيعي بحق واحد من اهم التوزيعات الاحتمالية الشائعة الاستخدام في النظرية الاحصائية وتطبيقاتها ، بسبب ان اغلب الظواهر الطبيعية تتبع هذا التوزيع . فاستخدامات هذا التوزيع تدخل في كافة الحقول والميادين كالزراعة والصناعة والطب والاجتماع وغيرها . ويعد هذا التوزيع القاعدة الاساس لموضوع الرقابة على جوده الانتاج Quality control ذا الاهمية الكبيرة في الصناعة . كذلك تبرز أهمية هذا التوزيع من خلال « نظرية الغاية المركزية central limit theorem (انظر الفقرة ٨ ـ ٢ ـ ٣) التي تثبت ان كافة التوزيعات الاحتمالية متقطعة كانت ام مستمرة يتقارب توزيعها (وفق شروط معينة) من التوزيع الطبيعي . إضف الى ذلك فان كل « توزيعات المعاينة Sampling dist » (لاحظ الفصل التاسع) يستند اشتقاقها للتوزيع الطبيعي. واستناداً لهذه الميزات اعتبر علماء الاحصاء هذا التوزيع اهم التوزيعات الاحتمالية على الاطلاق ويعتبر العالم الرياضي الانكليزي De – Moivre . او من اشتق دالة هذا التوزيع كحالة تقاربية من توزيع ثنائي الحدين وكان ذلك عام ١٧٣٣ . وفيما يلي تعريف لهذا التوزيع : يقال ان المتغير العشوائي X يتوزع كتوزيع طبيعي اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

$$f(x, a, b) = \frac{1}{\sqrt{2\pi b}}\, e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2}, \quad -\infty < x < \infty$$

حيث a,b تمثلان معلمتي التوزيع بحيث ان $-\infty < a < \infty$, $b > 0$ وسوف نلاحظ في فقرة لاحقة ان a هي متوسط هذا التوزيع وان b تمثل انحرافه المعياري .
والشكل (٦ ـ ٤) يوضح مخطط دالة هذا التوزيع .



الشكل (٦ ـ ٤) ، مخطط لدالة توزيع طبيعي .

ويمكن اثبات ان اجمـالي المسـاحة تحت منحنى الدالة $f(x)$ مسـاوية للواحد دلالة على كون $f(x)$ دالة كثافة احتمالية وكالآتي :

بفرض ان A تمثل المساحة تحت منحنى الدالة $f(x)$ اي ان

$$A = \int_{-\infty}^{\infty} f(x)\,dx = \frac{1}{\sqrt{2\pi}\ b} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx$$

بفرض ان

$$Z = \frac{x-a}{b} \quad \therefore x = bz + a \rightarrow dx = bdz\,,\ -\infty < Z < \infty$$

اذن

$$A = \frac{1}{\sqrt{2\pi}\ b} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}\, bdz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}\, dz$$

ان المطلوب برهانه هو ان $A = 1$ وهذا مكافيء لبرهان ان $A^2 = 1$ طالما ان $A > 0$

اذن :

$$A^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} z^2} dz . \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} y^2} dy$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 + y^2)} dz\, dy$$

ان حل التكامل المزدوج اعلاه، بهذا الشكل غير ممكن مما يستوجب الامر اجراء تحويل نحو الاحداثيات القطبية Polar coordinates من خلال التحويل الزاوي التالي :

$$\text{Sin}\, \theta = \frac{Z}{r} \rightarrow Z = r \sin \theta$$

$$\text{Cos}\, \theta = \frac{Y}{r} \rightarrow Y = r \,\text{Cos}\, \theta$$

$. r > 0 , 0 < \theta < 2\pi$

كما وان معامل التحويل من Z , Y الى r , θ معرف بالقيمة المطلقة لمحدد مصفوفة من مرتبة 2×2 عناصرها تمثل مشتقات جزئية للمتغيرين Z , Y نسبة الى r , θ. فاذا رمزنا لهذا المعامل بالرمز J فان :

$$J = \begin{vmatrix} \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \sin\theta & r\cos\theta \\ \cos\theta & -r\sin\theta \end{vmatrix} = -r(\sin^2\theta + \cos^2\theta)$$

$$= -r \qquad \therefore |J| = r$$

فاذن

$$A^2 = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} e^{-\frac{1}{2}(r^2 \cdot \sin^2\theta + r^2 \cdot \cos^2\theta)} \, |J| \, d\theta \, dr$$

$$= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} r e^{-\frac{1}{2}r^2} \, d\theta \, dr$$

$$= \frac{1}{2\pi} \int_0^\infty \left[\int_0^{2\pi} d\theta \right] r \, e^{-\frac{1}{2}r^2} \, dr$$

$$= \int_0^\infty r e^{-\frac{1}{2}r^2} \, dr = -\left[e^{-\frac{1}{2}r^2} \right]_0^\infty = 1 \qquad \therefore A = 1$$

ومما تقدم نستنتج ان

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx = \sqrt{2\pi} \, b,$$

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2\pi}$$

٦ ـ ٢ ـ ١ : الوسط والتباين في التوزيع الطبيعي .

سبق وان ذكرنا ان الوسط في التوزيع الطبيعي هو قيمة المعلمة a وان الانحراف المعياري هو قيمة المعلمة b وهذا يعني ان $\sigma_x^2 = b^2$.

البرهان :

$$EX = \frac{1}{\sqrt{2\pi} \, b} \int_{-\infty}^{\infty} x \cdot e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (bz + a) e^{-\frac{1}{2}z^2} dz, \; z = \frac{x-a}{b}$$

$$= \frac{b}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2} z^2} dz + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} z^2} dz$$

لكن

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2} z^2} dz = \sqrt{2\pi}$$

اذن

$$EX = \frac{b}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2} z^2} dz + a$$

كذلك فان

$$\int_{-\infty}^{\infty} z e^{-\frac{1}{2} z^2} dz = -[e^{-\frac{1}{2} z^2}]_{-\infty}^{\infty} = 0$$

عليه فان

$$EX = \mu_x = a$$

كذلك فان

$$\sigma_x^2 = EX^2 - (EX)^2$$

لكن

$$EX^2 = \frac{1}{\sqrt{2\pi}\, b} \int_{-\infty}^{\infty} x^2 e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (bz + a)^2 e^{-\frac{1}{2} z^2} dz$$

$$= \frac{b^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{1}{2} z^2} dz + \frac{2ab}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2} z^2} dz$$

$$+ \frac{a^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} z^2} dz$$

$$= \frac{b^2}{\sqrt{2\pi}} \int_{\infty}^{\infty} z^2 e^{-\frac{1}{2} z^2} dz + a^2$$

وباستخدام التكامل بطريقة التجزئة من خلال الفرض ان $u = z$،

$dv = z e^{-\frac{1}{2} z^2} dz$ يمكن اثبات ان $\int_{-\infty}^{\infty} z^2 e^{-\frac{1}{2} z^2} dz = \sqrt{2\pi}$

فاذن

$$EX^2 = b^2 + a^2$$

عليه فان :

$$\sigma_x^2 = b^2 + a^2 - a^2 = b^2 \qquad \therefore b = \sigma_x$$

وكما هو متداول في اغلب الأدبيات الاحصائية فانه يصار الى كتابة دالة الكثافة الاحتمالية لهذا التوزيع بالشكل :

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\ \sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} : -\infty < x < \infty$$

$$-\infty < \mu < \infty$$

$$\sigma > 0$$

وبالرموز فان $X \sim N(\mu, \sigma^2)$ اي ان المتغير العشوائي X يتوزع كتوزيع طبيعي بوسط μ وتباين σ^2 .

٦ ـ ٢ ـ ٢ : الدالة المولدة للعزوم حول نقطة الاصل .

ان الدالة المولدة لعزوم $N(\mu, \sigma^2)$ حول نقطة الاصل هي

$$M_X(t) = e^{\mu t + \frac{1}{2} t^2 \sigma^2}$$

البرهان :

$$M_X(t) = Ee^{tX} = \frac{1}{\sqrt{2\pi}\ \sigma} \int_{-\infty}^{\infty} e^{tx} . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma Z + \mu)}\ e^{-\frac{1}{2} Z^2}\ dz\ ,\ Z = \frac{x-\mu}{\sigma}$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(Z^2 - 2t\sigma Z)} dz$$

و باكمال المربع داخل القوس من خلال اضافة وطرح المقدار $t^2\sigma^2$ نحصل على :

$$M_X(t) = e^{\mu t + \frac{1}{2} t^2 \sigma^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(Z - t\sigma)^2} dz$$

لكن قيمة التكامل الاخير مُساوية للواحد وكأن $Z \sim N(t\sigma, 1)$. فاذن

$$M_X(t) = e^{\mu t + \frac{1}{2} t^2 \sigma^2}$$

ويتضح من هذه الدالة مايلي :

$1- M_X(0) = 1$

$2- K_X(t) = \ln M_X(t) = \mu t + \frac{1}{2} t^2 \sigma^2$

وان

$$K_x'(0) = \mu, K_x''(0) = \sigma^2, K_x^{(r)}(0) = 0, r = 3, 4, ...$$

كذلك يمكن اثبات ان الدالة المميزة لهذا التوزيع وبنفس الاسلوب اعلاه هي :

$$\phi(t) = e^{\mu i t + \frac{1}{2}(it)^2\sigma^2}$$
$$= e^{\mu i t - \frac{1}{2}t^2\sigma^2}$$

٦ ـ ٢ ـ ٣ : الدالة المولدة للعزوم المركزية

ان الدالة المولدة للعزوم المركزية لتوزيع $N(\mu, \sigma^2)$ هي :

$$M_{(x-\mu)}(t) = Ee^{t(X-\mu)} = e^{-\mu t}.M_x(t)$$
$$= e^{-\mu t}.e^{\mu t + \frac{1}{2}t^2\sigma^2} = e^{\frac{1}{2}t^2\sigma^2}$$

ويتضح من هذه الدالة ان .

$$M_{(x-\mu)}(t) = e^{\frac{1}{2}t^2\sigma^2} = \sum_{r=0}^{\infty} \frac{\left(\frac{1}{2}t^2\sigma^2\right)^r}{r!}$$

$$= \sum_{r=0}^{\infty} \frac{t^{2r}\sigma^{2r}}{2^r.r!} = \sum_{r=0}^{\infty} \frac{t^{2r}}{(2r)!} . \frac{(2r)!\sigma^{2r}}{2^r.r!}$$

وكما هو معلوم فان العزم المركزي ذو مرتبة r هو $M_{(x-\mu)}^{(r)}(0)$ او انه معامل $\frac{t^{2r}}{(2r)!}$ اي $\frac{(2r)!\sigma^{2r}}{2^r.r!}$ وهذا يعني ان

$$E(X-\mu)^{2r} = \frac{(2r)!\sigma^{2r}}{2^r.r!}$$

$$= \frac{\sigma^{2r}}{2^r.r!}.[2r(2r-1)(2r-2)...4\cdot3\cdot2\cdot1]$$

$$= \frac{\sigma^{2r}}{2^r . r!} [1\cdot3\cdot5 \ldots (2r - 1)] [2\cdot4\cdot6 \ldots (2r-2)(2r)]$$

$$= \frac{\sigma^{2r}}{2^r . r!} [1\cdot3\cdot5 \cdots (2r - 1)]. [2^r . r!]$$

$$= \sigma^{2r} . \prod_{k=1} (2K - 1)$$

وبشكل خاص فان

$$E(X-\mu)^{2r}\Big]_{r=1} = \sigma^2 , E(X-\mu)^{2r}\Big]_{r=2} = 3\sigma^4 , E(X-\mu)^{2r}\Big]_{r=3} = 15\sigma^6 , \ldots$$

وان

$$E(X-\mu)^{2r-1} = 0 ; r = 1, 2, 3, \ldots$$

وهذا يعني ان العزوم المركزية ذات مراتب فردية مساوية للصفر دائماً ونستنتج مما تقدم ان :

معامل الالتواء في هذا التوزيع هو $S_K = \frac{\mu_3^2}{\sqrt{\mu_2^3}} = 0$. حيث μ_3 , μ_2 هما على التوالي العزم المركزي الثاني والثالث . وحيث ان $S_K = 0$ فذلك يعني ان منحنى دالة هذا التوزيع متماثل .

٦ ـ ٢ ـ ٤ : المنوال والوسيط في التوزيع الطبيعي

نظراً لخاصية التماثل في منحنى دالة التوزيع الطبيعي فانه يمكن اثبات ان وسط هذا التوزيع مساوٍ لمنواله وفي الوقت ذاته مساوٍ لوسيطه وكما يلي :
ان المنوال يمثل قيمة x الناتجة من حل المعادلة التفاضلية $f'(x)=0$ بشرط ان $f''(x) < 0$. الان

$$f(x) = \frac{1}{\sqrt{2\pi}\ \sigma} . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

فاذن

$$\ln f(x) = \ln\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right) - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$\frac{d\ln f(x)}{dx} = \frac{f'(x)}{f(x)} = -\frac{1}{\sigma^2}(x-\mu)$$

او ان

$$f'(x) = -f(x) \cdot \frac{(x-\mu)}{\sigma^2}$$

وان

$$f''(x) = -\frac{1}{\sigma^2}\left[(x-\mu)f'(x) + f(x)\right]$$

$$= -\frac{f(x)}{\sigma^2}\left[1 - \left(\frac{x-\mu}{\sigma}\right)^2\right]$$

الان بجعل $f'(x) = 0$ فان

$$-f(x)\frac{(x-\mu)}{\sigma^2} = 0$$

اي ان

$$f(x)(x-\mu) = 0$$

وحيث ان $f(x) > 0$ فاذن $x - \mu = 0$ او $x = \mu$ ، وان

$$f''(x)\Big]_{x=\mu} = -\frac{f(x)}{\sigma^2} < 0$$

وهذا يعني ان $x = \mu$ يمثل المنوال الوحيد لهذا التوزيع . وان القيمة العظمى للدالة $f(x)$ هي

$$\operatorname{Max} f(x) = f(x)\Big]_{x=\mu} = \frac{1}{\sqrt{2\pi}\,\sigma}$$

ويلاحظ هنا ان قيمة σ هي التي تتحكم بالقيمة العظمى للدالة $f(x)$ فكلما كان تباين التوزيع عالٍ فذلك مؤشر انخفاض القيمة العظمى والعكس صحيح ايضاً .

كما وان الوسيط M ناتج من حل المعادلة التكاملية :

$$\int_{-\infty}^{M} f(x)\,dx = \frac{1}{2}$$

$$= \int_{-\infty}^{\mu} f(x)\,dx + \int_{\mu}^{M} f(x)\,dx = \frac{1}{2}$$

وحيث ان منحنى التوزيع متماثل عند $x = \mu$ (اي المنوال) فذلك يعني ان المساحة تحت منحنى الدالة $f(x)$ للفترة $(-\infty, \mu)$ تساوي $\frac{1}{2}$. فاذن

$\int_{\mu}^{M} f(x)\,dx = 0$ وهذا يعني ان $M = \mu$ ويلاحظ هنا ان :

الوسط = الوسيط = المنوال = μ

٦ ـ ٢ ـ ٥ : نقاط الانقلاب والشكل العام لمنحنى دالة التوزيع الطبيعي :

لمنحنى دالة التوزيع الطبيعي نقطتا انقلاب تقعان على بعد متساوي الى يمين ويسار المنوال . ولاحظنا في الفقرة (٦ ـ ٢ ـ ٤) ان :

$$f''(x) = -\frac{f(x)}{\sigma^2}\left[1 - \left(\frac{x-\mu}{\sigma}\right)^2\right]$$

وكما هو معلوم فان نقاط الانقلاب لمنحنى دالة مستمرة ناتجة من حل المعادلة التفاضلية $f''(x) = 0$ بشرط ان $f'''(x) \neq 0$. فاذن بوضع $f''(x) = 0$ نحصل على

$$-\frac{f(x)}{\sigma^2}\left[1 - \left(\frac{x-\mu}{\sigma}\right)^2\right] = 0$$

او ان

$$f(x)\left[1 - \left(\frac{x-\mu}{\sigma}\right)^2\right] = 0$$

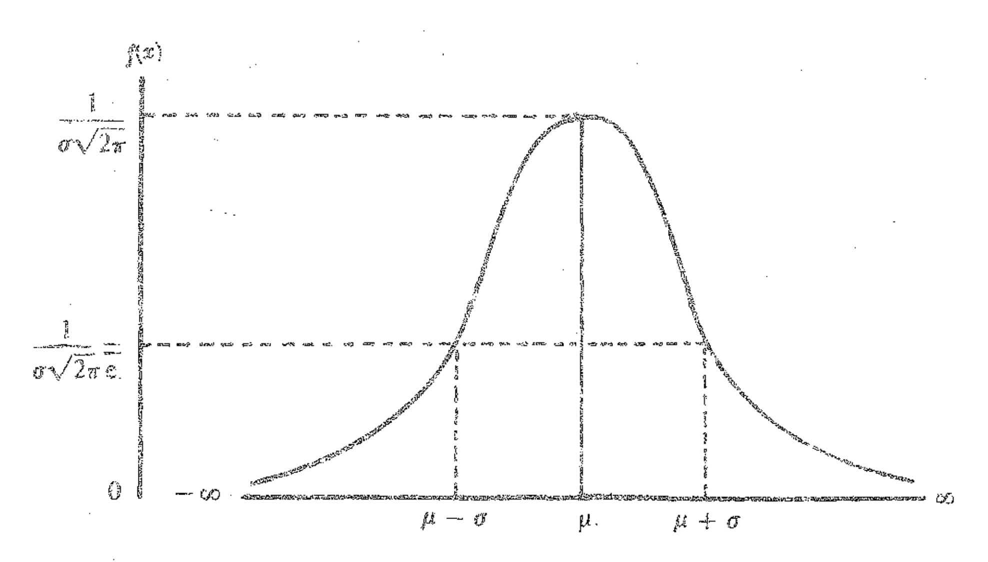
وحيث ان $f(x) > 0$ فاذن

$$1 - \left(\frac{x-\mu}{\sigma}\right)^2 = 0 \rightarrow (x-\mu)^2 = \sigma^2$$

او ان

$$x - \mu = \pm \sigma \rightarrow x = \mu \pm \sigma$$

ويمكن اثبات ان $f'''(x)$ عند $x = \mu - \sigma$ او $x = \mu + \sigma$ غير مساوية للصفر . وهذا يعني ان نقطتا الانقلاب في منحنى دالة هذا التوزيع هما $x = \mu + \sigma , x = \mu - \sigma$ لاحظ ان هاتين النقطتين تقعان على بعد ثابت قدره نحو يمين ويسار المنوال . عليه فان الشكل العام لمخطط دالة $N(\mu, \sigma^2)$ هو الموضح في الشكل (٦ ـ ٥) :

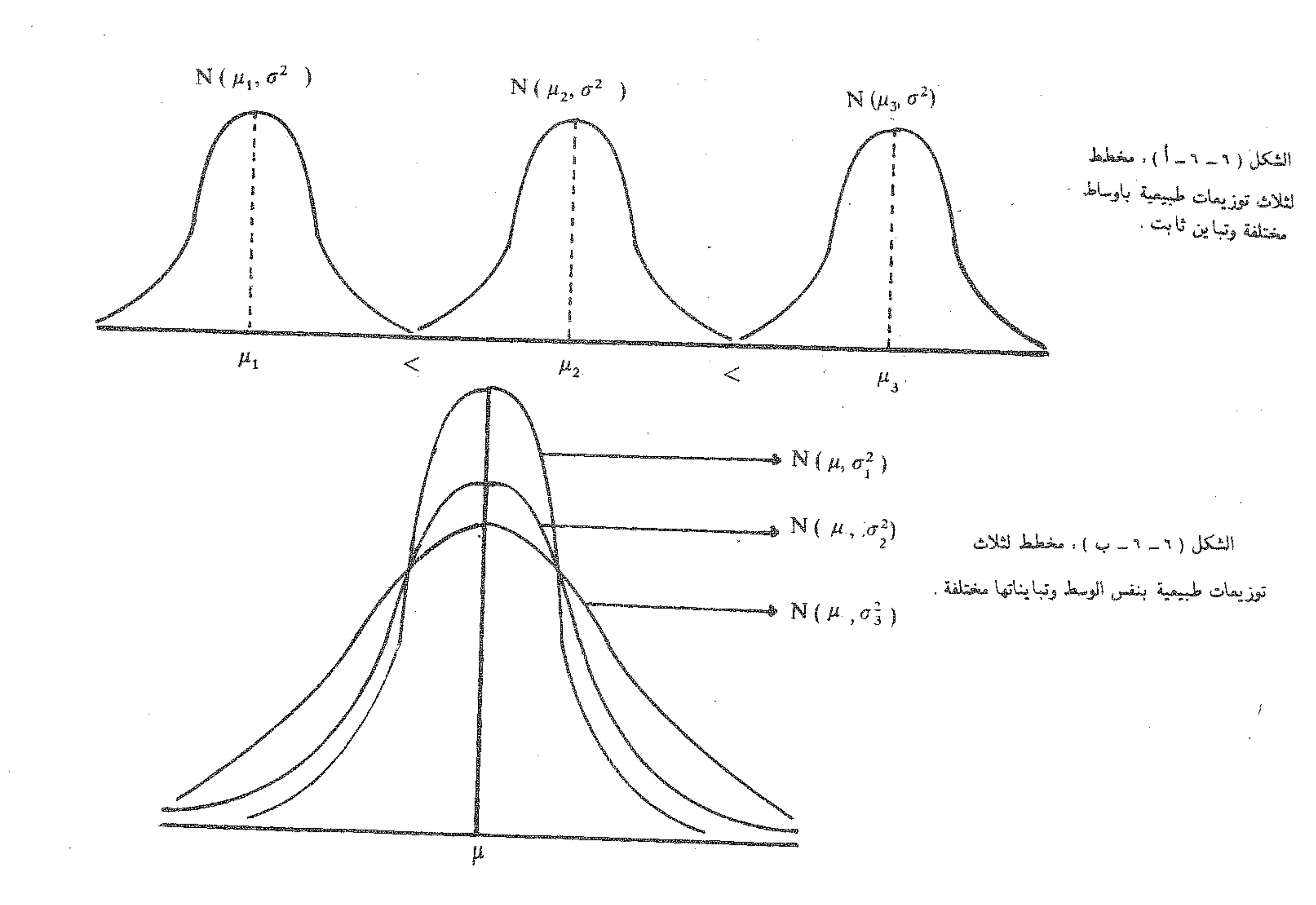


الشكل (٦ ـ ٥) ، مخطط لدالة توزيع $N(\mu, \sigma^2)$

وان قيمة الدالة $f(x)$ عند نقطتي الانقلاب هي :

$$f(x)\Big]_{x=\mu-\sigma} = f(x)\Big]_{x=\mu+\sigma} = \frac{1}{\sqrt{2\pi e}\ \sigma}$$

ان قيمة μ تحدد موقع التوزيع على المحور السيني في حين ان σ^2 تبين مقدار تفلطح منحنى دالة هذا التوزيع . فكلما كانت σ^2 كبيرة فان ذلك مؤشر على تفلطح منحنى هذا التوزيع (اي ان تشتت قيم X عال) في حين كلما كانت σ^2 صغيرة فان ذلك مؤشر على تدبدب متحناه (اي ان قيم المتغير X اكثر تجانساً) . والشكل (٦ ـ ٦ ـ أ) و (٦ ـ ٦ ـ ب) يوضحان ماتقدم .



الشكل (٦ ـ ٦ ـ أ) : مخطط لثلاث توزيعات طبيعية باوساط مختلفة وتباين ثابت .

الشكل (٦ ـ ٦ ـ ب) : مخطط لثلاث توزيعات طبيعية بنفس الوسط وتبايناتها مختلفة .

٦ ـ ٢ ـ ٦ : التوزيع الاحتمالي لتركيب خطي .

لتكن $X_1, X_2, \ldots, X_n$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim N(\mu_i, \sigma_i^2)$ وان $a_1, a_2, \ldots, a_n$ ثوابت حقيقية . ليكن $Y = \sum_{i=1}^{n} a_i X_i$ يمثل تركيباً خطياً Linear combination بدلالة هذه المتغيرات عندئذ فان

$$Y \sim N\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$$

البرهان : افرض ان $M_Y(t)$ موجودة . فاذن

$$M_Y(t) = Ee^{tY} = Ee^{t\sum_{i=1}^{n} a_i X_i}$$

$$= \prod_{i=1}^{n} Ee^{t_i^* X_i}, \; t_i^* = ta_i$$

وحيث ان $X_i \sim N(\mu_i, \sigma_i^2)$ فاذن

$$Ee^{t_i^* X_i} = M_{X_i}(t_i^*) = e^{\mu_i t_i^* + \frac{1}{2} t_i^{*2}\sigma_i^2}$$

عليه فان

$$M_Y(t) = \prod_{i=1}^{n} e^{\mu_i t_i^* + \frac{1}{2} t_i^{*2}\sigma_i^2}$$

$$= e^{t\sum_{i=1}^{n} a_i\mu_i + \frac{1}{2}t^2 \sum_{i=1}^{n} a_i^2\sigma_i^2}$$

والصيغة الاخيرة هي الدالة المولدة لعزوم توزيع طبيعي بوسط $\sum_{i=1}^{n} a_i\mu_i$ وتبـاين $\sum_{i=1}^{n} a_i^2\sigma_i^2$ فاذن $\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$.

وبشكل خاص واستناداً لهذه النظرية فانه :

١ ـ اذا كانت $\forall\, i = 1, 2, \ldots, n \quad a_i = 1$ فان :

$$\sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i , \sum_{i=1}^{n} \sigma_i^2 \right)$$

٢ ـ اذا كانت $n = 2 , a_1 = 1 , a_2 = -1$ فان

$$X_1 - X_2 \sim N(\mu_1 - \mu_2 , \sigma_1^2 + \sigma_2^2)$$

٣ ـ اذا كانت $\forall\, i \quad a_i = \frac{1}{n} , \mu_i = \mu , \sigma_i^2 = \sigma^2$ فان

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i \sim N\left(\mu , \frac{\sigma^2}{n}\right)$$

وهذا يعني ان متوسط قياسات عينة عشوائية مختارة من مجتمع ذي توزيع $N(\mu, \sigma^2)$ قوامها n مفردة يتوزع طبيعياً بوسط μ وتباين $\frac{\sigma^2}{n}$.

٦ ـ ٢ ـ ٧ ـ التوزيع الطبيعي المعياري
Standard normal distribution

لاحظنا من الفقرات السابقة ان هنالك عدداً غير منتة (عائلة) من التوزيعات الطبيعية التي يمكن تحديد اي منها من خلال معرفة قيمة μ , σ^2 . وهذا يعني ان لكل عضو من هذه العائلة دالة توزيعية عن طريقها يمكن بناء جدول التراكمات الاحتمالية . الا ان هذا امر غير مجدٍ من الناحية العملية حيث انه يتطلب بناء جدول لكل توزيع منها مما يقتضي ذلك ايجاد شكل ثابت لهذا التوزيع واستناداً لهذا الشكل يمكن بناء هذا الجدول . ان هذا الشكل يسمى التوزيع الطبيعي المعياري او توزيع الدرجة المعيارية في $N(\mu, \sigma^2)$. وفيما يلي اشتقاق لهذا التوزيع .

بفرض ان $X \sim N(\mu, \sigma^2)$ وان $Z = \frac{X - \mu}{\sigma}$ وبفرض ان الدالة المولدة لعزوم Z موجودة فذلك يعني ان :

$$M_Z(t) = Ee^{tZ}$$

$$= Ee^{t\left(\frac{X-\mu}{\sigma}\right)} = e^{-t \cdot \frac{\mu}{\sigma}} \cdot M_X\left(\frac{t}{\sigma}\right)$$

$$= e^{-t\frac{\mu}{\sigma}} \cdot e^{t\frac{\mu}{\sigma} + \frac{1}{2}\left(\frac{t}{\sigma}\right)^2 \cdot \sigma^2}$$

$$= e^{\frac{1}{2}t^2}$$

والصيغة الاخيرة هي الدالة المولدة لعزوم توزيع طبيعي بوسط صفر وتباين واحد . وهذا يعني ان

$$Z \sim N(0, 1)$$

فاذن

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad -\infty < Z < \infty$$

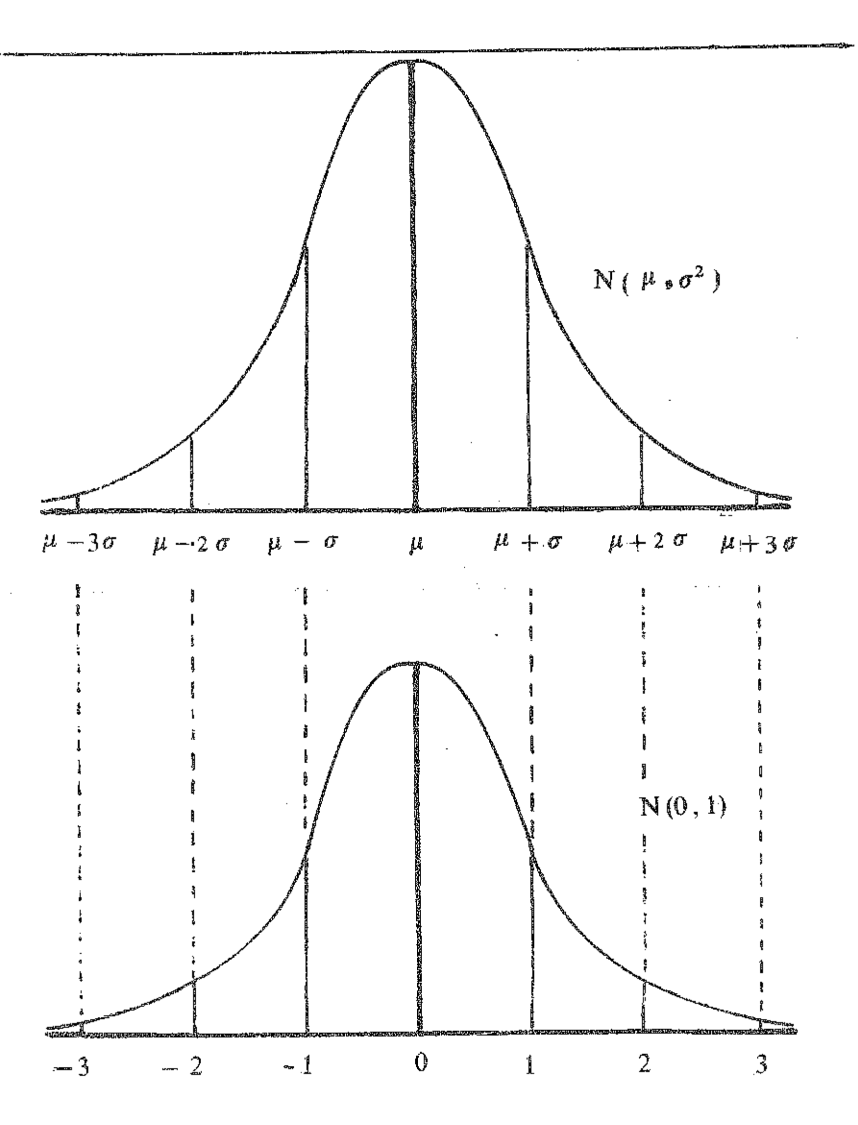
ومما تقدم نلاحظ انه يمكن تحويل اي توزيع طبيعي مهما كان وسطه وتباينه الى التوزيع الطبيعي المعياري وفق التحويل $Z = \frac{X - \mu}{\sigma}$. فمثلاً اذا كان

$$X \sim N(10, 4) \rightarrow Z = \frac{X - 10}{2} \sim N(0, 1)$$

$$X \sim N(100, 64) \rightarrow Z = \frac{X - 100}{8} \sim N(0, 1)$$

$$X \sim N(-50, 9) \rightarrow Z = \frac{X + 50}{3} \sim N(0, 1)$$

وهذا يعني ثبات التوزيع الطبيعي عند وسط قدرة صفر وتباين مقداره واحد الامر الذي يمكننا من بناء جدول خاص بهذا التوزيع. علماً ان خصائص $N(0,1)$ هي نفس خصائص $N(\mu,\sigma^2)$ بمجرد التعويض عن $\mu = 0 , \sigma^2 = 1$ والشكل (٦ ـ ٧) يوضح مقارنة بين $N(0,1), N(\mu,\sigma^2)$ المشتق منه.



الشكل (٦ ـ ٧) : مقارنة بين $N(0,1), N(\mu,\sigma^2)$

٦ ـ ٢ ـ ٨ : الدالة التوزيعية Distribution function

استناداً لما تم توضيحه في الفقرة السابقة يمكن تعريف الدالة التوزيعية وبناء جداول خاصة بالتوزيع الطبيعي كما يلي : ان

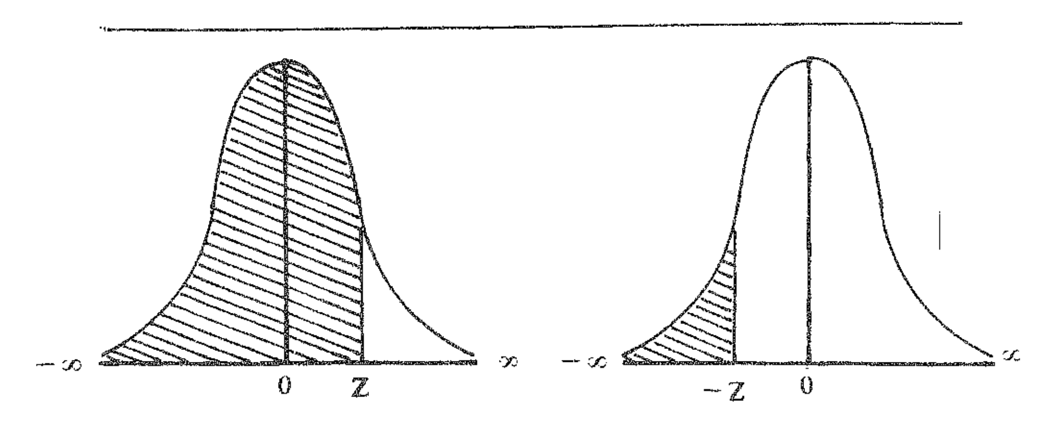
$$F(x) = P_r(X \leq x) = \int_{-\infty}^{x} f(u)\,du \ , u \sim N(\mu, \sigma^2)$$

$$= P_r\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma} \right) = P_r(Z \leq z) = F(z)$$

$$= \int_{-\infty}^{z} f(v)\,dv \ , v \sim N(0, 1)$$

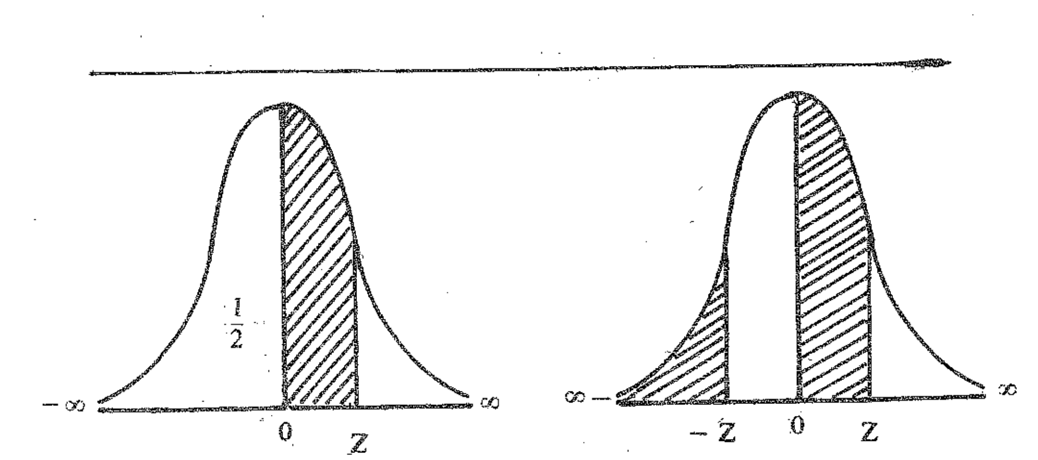
وعلى اساس الدالة $F(z)$ تم بناء جداول التوزيع الطبيعي التي تبين الاحتمال المتراكم لغاية قيمة معطاة الى z . هذه الجداول على نوعين رئيسين هما :

أ ـ جداول مبنية على اساس التكامل على الفترة $(-\infty, z)$ (لاحظ الجدول ٤ ملحق ب). هذا النوع من الجداول هو الافضل والاسهل تداولاً في النواحي العملية وفكرة بناء هذه الجداول هي حساب المساحة تحت منحنى دالة $N(0,1)$ للفترة $(-\infty, z)$ وكما هو موضح في الشكل (٦ ـ ٨) وبفرض ان $z > 0$:



الشكل (٦ ـ ٨) : توضيح لحساب المساحة تحت منحنى دالة $N(0,1)$

ب ـ جداول مبنية على اساس الاستفادة من خاصية التماثل في هذا التوزيع . وفكرة بناء هذه الجداول هي حساب المساحة تحت منحنى دالة $N(0,1)$ للفترة $(0, z)$ ومن ثم يضاف $\frac{1}{2}$ للناتج باعتبار ان المساحة للفترة $(-\infty, 0)$ مساوية للنصف . اما اذا كانت z سالبة عندئذٍ يتم حساب $F(-z)$ على اساس طرح المساحة للفترة $(0, z)$ من $\frac{1}{2}$ باعتبار ان المساحة للفترة $(-z, 0)$ تساوي المساحة للفترة $(0, z)$ كما هو موضح بالشكل (٦ ـ ٩) :



الشكل (٦ ـ ٩) : توضيح لحساب المساحة تحت منحنى دالة $N(0,1)$

وبشكل عام فان :

$$1 - P_r(Z < 0) = P_r(Z > 0) = 0.5$$
$$2 - P_r(Z < -z) = 1 - P_r(Z < z), z > 0$$
$$3 - P_r(z_1 < Z < z_2) = P_r(Z < z_2) - P_r(Z < z_1) = F(z_2) - F(z_1)$$

واذا كان $X \sim N(\mu, \sigma^2)$ فان :

$$1 - P_r(\mu - \sigma < X < \mu + \sigma) = P_r(-1 < Z < 1) = F(1) - F(-1) = 0.8413 - 0.1587 = 0.6826$$

$$2 - P_r(\mu - 2\sigma < X < \mu + 2\sigma) = F(2) - F(-2)$$
$$= 0{\cdot}9772 - 0{\cdot}0228 = 0{\cdot}9544$$
$$3 - P_r(\mu - 3\sigma < X < \mu + 3\sigma) = F(3) - F(-3)$$
$$= 0{\cdot}9987 - 0{\cdot}0013 = 0{\cdot}9974$$

وكما هو موضح في الشكل (٦ ــ ٧) . مع ملاحظة ان كتب وادبيات الاحصاء تتباين فيما بينها من حيث الرمز المخصص للدالة التوزيعية فالبعض يخصص الرمز $F(.)$ والبعض الاخر يخصص الرمز $\Phi(\cdot)$ واخرون الرمز $N(\cdot)$ واياً كان الرمز فان المضمون هو نفسه اي « الدالة التوزيعية في توزيع $N(0, 1)$ » .

٦ ــ ٢ ــ ٩ : اسلوب بناء جداول التوزيع الطبيعي .

هنالك طرق عديدة تم من خلالها بناء جداول التوزيع الطبيعي تباينت فيما بينها من حيث الدقة في حساب $F(Z)$ نذكر منها اربعة طرق فقط وهي :

١ ــ استخدام سلسلة تايلر في حساب قيم $F(Z)$

يعد هذا الاسلوب اقدم طريقة لحساب قيم $F(z)$ حيث يستند على فك المقدار $e^{-\frac{1}{2}z^2}$ باستخدام سلسلة تايلر وكما يلي :

$$F(z) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{1}{2}u^2} du$$

$$= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{2}u^2\right)^k}{k!} du$$

$$= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z \left(1 - \frac{u^2}{2(1!)} + \frac{u^4}{4(2!)} - \frac{u^6}{8(3!)} + \frac{u^8}{16(4!)} - \dots\right) du$$

ويتم اجراء التكامل لكل حد من حدود القوس على الفترة $(0, z)$. ويلاحظ ان درجة الدقة المطلوبة في حساب $F(Z)$ تعتمد على عدد الحدود التي يتوقف عندها فك السلسلة . ولغرض التوضيح فقط سنتوقف عند الحد الاخير الموضح داخل القوس . وبفرض ان $z = 1$ اي نرغب في حساب $F(1)$ فان

$$F(1) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left(Z - \frac{Z^3}{6} + \frac{Z^5}{40} - \frac{Z^7}{336} + \right.$$

$$\left. + \frac{Z^9}{3456} \right) \Big]_0^1$$

$$= 0{\cdot}8413534$$

وهي تقريبا نفس قيمة $F(1)$ الواردة في الجدول (٤) ملحق (ب) . وعلى هذا الاساس يتم حساب قيمة $F(Z)$ بعد تخصيص قيمة الى Z .

٢ ـ الاسلوب المقترح من قبل پو ليا Po'lya عام ١٩٤٥ .

تمكن پوليا من التوصل للصيغة التالية لحساب $F(Z)$ والتي تعد ادق من الصيغة الاولى (عندما يكون عدد حدود المفكوك قليل) وهي :

$$F(Z) = \frac{1}{2} \left[1 + \left(1 - e^{-\frac{2z^2}{\pi}} \right)^{\frac{1}{2}} \right]$$

ولغرض المقارنة فان $F(1)$ وفق هذه الصيغة مساوية الى $0{\cdot}8431188$. ان اعظم خطأ يحصل وفق صيغة پوليا هو 0.003 عندما $Z = 1{\cdot}6$

٣ ـ الاسلوب المقترح من قبل كادويل Cadwell عام ١٩٥١ .

اقترح كادويل الصيغة التالية لحساب $F(Z)$ والتي تعد اكثر دقة من الطريقة السابقة ، وهي :

$$F(Z) = \frac{1}{2}\left[1 + (1 - e^{k})^{\frac{1}{2}}\right]$$

حيث

$$K = -\frac{2Z^2}{\pi}\left(1 + \frac{(\pi - 3)Z^2}{3\pi}\right)$$

فمثلاً $F(1)$ وفق صيغة كادويل مساوية الى 0.8449486 ،ان اعظم خطأ يحصل في حساب $F(Z)$ وفق هذه الصيغة هو 0.0007 عندما $Z = 2.5$ ، كذلك بين كادويل انه عند اضافة المقدار $(-0.0005\,Z^6 + 0.00002\,Z^8)$ الى K فان هذا الاجراء يختزل هذا الخطأ الى 0.00005 .

٤ ـ الاسلوب المقترح من قبل موران Moran عام ١٩٨٠ .

تمكن موران من تطوير صيغة تقريبية للدالة $F(Z)$ يمكن اعتبارها ادق الصيغ المقترحة سابقاً حيث ان اعظم خطأ يحصل في حساب $F(Z)$ هو 10^{-9}. هذه الصيغة هي :

$$F(Z) = \frac{1}{2} + \frac{1}{\pi}\left[\frac{Z}{3\sqrt{2}} + \sum_{n=1}^{\infty}\frac{e^{-\frac{n^2}{9}}}{n}\,\mathrm{Sin}\,\frac{nZ\sqrt{2}}{3}\right]$$

وقد لاحظ موران (وللاغراض التطبيقية) انه يمكن اجراء عملية الجمع لغاية n = 12 بحيث ان ذلك يجعل قيمة اعظم خطأ ممكن في حساب $F(Z)$ هو 10^{-9} ان قيمة $F(1)$ وفق صيغة موران هي 0.8414043

٦ ـ ٢ ـ ١٠ : التوزيع الطبيعي المبتور

Truncated normal distribution

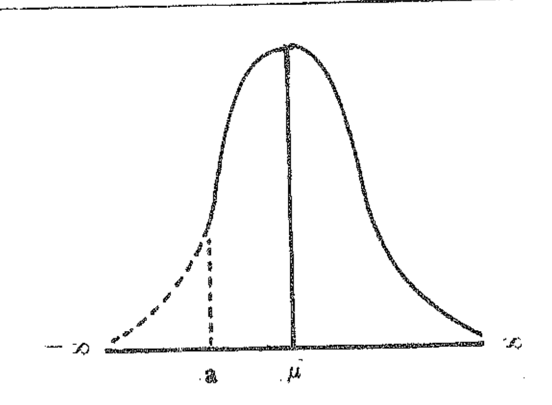
يتطلب الامر في بعض الاحيان عمل بتر في التوزيع الطبيعي من خلال حذف جزء من القيم الممكنة لهذا التوزيع وتعريف الدالة $f(x)$ على بقية القيم الإخرى . فعلى فرض ان $X \sim N(\mu, \sigma^2)$ فان بعض اشكال البتر الممكنة في هذا التوزيع هي :

١ ـ البتر من جهة اليسار :

ليكن a ثابتاً حقيقياً بحيث ان $a < \mu$ وتطلب الامر تعريف الدالة $f(x)$ على الفترة (a, ∞) فاذن ذلك يتم من خلال مايلي :

$$f(x \mid x > a) = \frac{f(x)}{1 - F(a)}\ ; a < x < \infty$$

وكما هو موضح في الشكل (٦ ـ ١٠) :



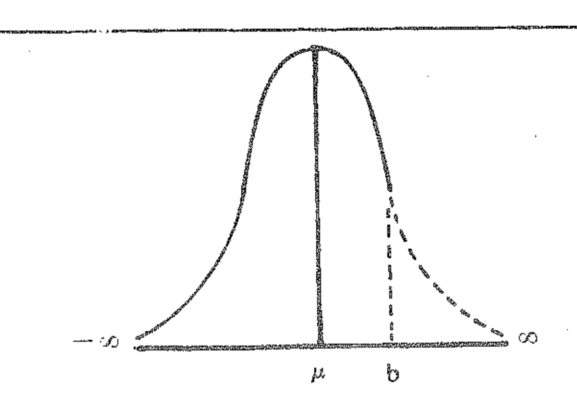
الشكل (٦ ـ ١٠) : بتر في التوزيع الطبيعي من جهة اليسار .

٢ ـ البتر من جهة اليمين :

افرض ان b ثابت حقيقي بحيث ان $\mu < b$ وتطلب الامر تعريف الدالة $f(x)$ على الفترة $(-\infty , b)$ فان ذلك يتم من خلال ما يلي :

$$f(x \mid x < b) = \frac{f(x)}{F(b)} , -\infty < x < b$$

والشكل (٦ ـ ١١) يوضح ذلك :



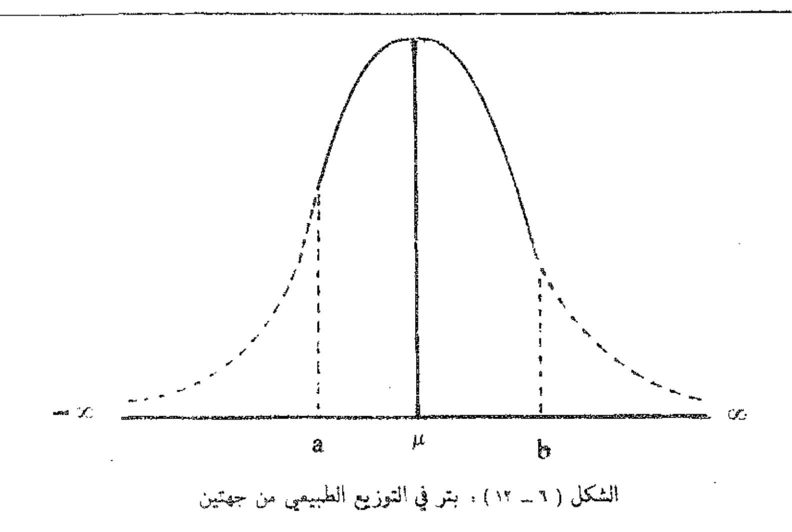
الشكل (٦ ـ ١١) : بتر في التوزيع الطبيعي من جهة اليمين

٣ ـ البتر من جهتين :

ليكن a , b ثابتان حقيقيان بحيث ان $a < \mu < b$ وتطلب الامر تعريف الدالة $f(x)$ على الفترة (a , b) فان ذلك يتم وفق ما يلي :

$$f(x) = \frac{f(x)}{F(b) - F(a)} , a < x < b$$

وكما موضح في الشكل (٦ ـ ١٢) :



الشكل (٦ ـ ١٢) : بتر في التوزيع الطبيعي من جهتين

علماً ان هنالك اشكال اخرى للبتر حسبما تقتضيه الحالة تحت الدراسة . كذلك فان عمليات البتر بشكل عام تؤثر في مؤشرات التوزيع كالوسط والتباين وغيرها من الامور ذات العلاقة بالتوزيع فمثلاً الوسط للتوزيع الناتج بعد اجراء عملية البتر من جهتين هو :

$$E(X|a < x < b) = \frac{1}{F(b) - F(a)} \cdot \frac{1}{\sqrt{2\pi}\ \sigma} \int_a^b x . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{F(b) - F(a)} \cdot \frac{1}{\sqrt{2\pi}} \int_{a^*}^{b^*} (\sigma z + \mu) e^{-\frac{1}{2} z^2} dz$$

حيث

$$a^* = \frac{a - \mu}{\sigma}, b^* = \frac{b - \mu}{\sigma}$$

$$= \frac{1}{F(b) - F(a)} \left[\frac{\sigma}{\sqrt{2\pi}} \int_{a^*}^{b^*} Z e^{-\frac{1}{2} z^2} dz + \frac{\mu}{\sqrt{2\pi}} \int_{a^*}^{b^*} e^{-\frac{1}{2} z^2} dz \right]$$

لكن

$$\frac{1}{\sqrt{2\pi}} \int_{a^*}^{b^*} Z e^{-\frac{1}{2} z^2} dZ = - \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} \right]_{a^*}^{b^*}$$

$$= - [f(z)]_{a^*}^{b^*} = f(a^*) - f(b^*)$$

وان

$$\frac{1}{\sqrt{2\pi}} \int_{a^*}^{b^*} e^{-\frac{1}{2} z^2} dz = F(b^*) - F(a^*).$$

فاذن

$$E(X \mid a < x < b) = \frac{1}{F(b) - F(a)} [\sigma(f(a^*) - f(b^*) + \mu(F(b^*) - F(a^*)]$$

لكن

$$f(z = a^*) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} a^{*2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{a-\mu}{\sigma} \right)^2}$$

$$= \sigma f(x = a) = \sigma f(a)$$

وان

$$f(z = b^*) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} b^{*2}} = \frac{1}{\sqrt{2\pi}} {}^{-\frac{1}{2} \left(\frac{b-\mu}{\sigma} \right)^2}$$

$$= \sigma f(x = b) = \sigma f(b)$$

وان

$$F(a^*) = P_r(Z \leq a^*) = P_r \left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma} \right)$$

$$= P_r (X \le a) = F(a)$$

كذلك فان

$$F(b^*) = P_r(Z \le b^*) = P_r\left(\frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right)$$

وان

$$= P_r(X \le b) = F(b)$$

عليه فان

$$E(X|a < x < b) = \frac{1}{F(b) - F(a)} [\sigma^2 (f(a) - f(b)]$$

$$+ \mu (F(b) - F(a))]$$

$$= \mu + \frac{f(a) - f(b)}{F(b) - F(a)} . \sigma^2$$

كذلك يمكن اثبات ان الدالة المولدة لعزوم X بعد اجراء بتر في التوزيع من جهتين هي :

$$M_x(t) = \frac{F(b^* - \sigma t) - F(a^* - \sigma t)}{F(b) - F(a)}$$

ونترك برهنة ذلك للقاريء.

٩ ـ ٢ ـ ١١ : توزيع القيمة المطلقة لمتغيرذا توزيع طبيعي معياري

The distribution of absolute standard normal variate

ليكن X متغير عشوائي يتوزع وفق دالة توزيع $N(\mu, \sigma^2)$ وان $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$ عندئذ فان $V = |Z|$ سوف يتوزع كتوزيع طبيعي مبتور على الفترة $(0, \infty)$ وهذا يعني ان

$$f(V) = f(Z|Z \geq 0) = \frac{f(z)}{1 - F(0)}$$

وحيث ان $F(0) = \frac{1}{2}$ لذا فان

$$f(V) = 2f(Z) = 2 . \frac{1}{\sqrt{2\pi}} . e^{-\frac{1}{2}z^2}$$

$$= \sqrt{\frac{2}{\pi}} . e^{-\frac{1}{2}z^2} \; ; Z \geq 0$$

وبالرموز فان $V = |Z| \sim AN(0,1)$ ويسمى هذا التوزيع في بعض الاحيان بالتوزيع الطبيعي الموجب positive normal dist .

وفيما يلي بعض خصائص هذا التوزيع .

١ ـ ان الوسط والتباين للمتغير $V = |Z|$ هما $\mu_v = \sqrt{\frac{2}{\pi}}$, $\sigma^2_V = 1 - \frac{2}{\pi}$

البرهان :

$$\mu_v = \sqrt{\frac{2}{\pi}} \int_0^\infty |Z| e^{-\frac{1}{2}z^2} dz$$

وحيث ان $Z = |Z|$ في الفترة $(0, \infty)$ فاذن

$$\mu_v = \sqrt{\frac{2}{\pi}} \int_0^\infty Z e^{-\frac{1}{2}z^2} dz = \sqrt{\frac{2}{\pi}} [- e^{-\frac{1}{2}z^2}]_0^\infty$$

$$= \sqrt{\frac{2}{\pi}}$$

وهنا نلاحظ ان $\mu_v = E|Z|$ ماهو الا الانحراف المتوسط في توزيع $N(0,1)$

كذلك فان ؛

$$\sigma_V^2 = EV^2 - (EV)^2$$

لكن

$$EV^2 = \sqrt{\frac{2}{\pi}} \int_0^\infty Z^2 e^{-\frac{1}{2}z^2} dz$$

وباستخدام التكامل بالتجزئة من خلال الفرض $Z = u$, $Ze^{-\frac{1}{2}Z^2} dz = dv$

والاستفادة من الخاصية $\int_0^\infty e^{-\frac{1}{2}z^2} dz = \frac{1}{2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz$ يمكن ملاحظة ان

$EV^2 = 1$ فاذن

$$\sigma_V^2 = 1 - \frac{2}{\pi}$$

٢ ـ العزم ذو المرتبة r حول نقطة الاصل

ان العزم ذو المرتبة r حول نقطة الاصل في هذا التوزيع هو $EV^r = \dfrac{2^{r/2}\,\Gamma\left(\frac{r+1}{2}\right)}{\pi}$

البرهان :

$$EV^r = \sqrt{\frac{2}{\pi}} \int_0^\infty V^r e^{-\frac{1}{2}v^2} dv$$

وبفرض ان

$$g = \frac{V^2}{2} \quad \therefore V = \sqrt{2g} \rightarrow dv = \frac{dg}{\sqrt{2g}} ,\ 0 < g < \infty$$

فاذن

$$EV^r = \sqrt{\frac{2}{\pi}} \int_0^\infty (2g)^{\frac{r}{2}} \cdot e^{-g} \cdot (2g)^{-\frac{1}{2}} dg$$

$$= \frac{2^{\frac{r}{2}}}{\sqrt{\pi}} \int_0^\infty g^{\left(\frac{r+1}{2}\right)-1} \cdot e^{-g} dg$$

ان التكامل الاخير هو تكامل كاما (لاحظ الفقرة ٦ ـ ٤) بالمعلمتين $\beta = 1, \alpha = \frac{r+1}{2}$ وان قيمة هذا التكامل هي $\Gamma\left(\frac{r+1}{2}\right)$ عليه فان

$$EV^r = \frac{2^{r/2}\,\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}} \quad , r = 1, 2, 3, \ldots$$

علماً انه لاي عدد مثل n فان $\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)!$ وان $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ وسوف نبرهن ذلك في الفقرة (٦ ـ ٤) . لاحظ من هذه الصيغة انه عندما :

$r = 1$ فان $\Gamma(1) = 0! = 1$ ،

$$EV = \mu_V = \frac{2^{\frac{1}{2}}\,\Gamma(1)}{\sqrt{\pi}} = \sqrt{\frac{2}{\pi}}$$

$r = 2$ فان

$$EV^2 = \frac{2\Gamma\left(\frac{3}{2}\right)}{\sqrt{\pi}} = \frac{2 \cdot \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi}} = 1$$

٣ ـ الدالة المولدة للعزوم حول نقطة الاصل .

يمكن اثبات ان الدالة المولدة لعزوم المتغير $V = |Z|$ هي

$$M_V(t) = 2M_Z(t) \, . \, F(t)$$

حيث $M_Z(t)$ هي الدالة المولدة لعزوم توزيع $N(0,1)$ حول نقطة الاصل وان $F(t)$ تمثل الدالة التوزيعية لتوزيع $N(0,1)$ عند $Z = t$

البرهان :

$$M_V(t) = Ee^{tV} = \sqrt{\frac{2}{\pi}} \int_0^\infty e^{tV} . e^{-\frac{1}{2}V^2} dv$$

$$= \sqrt{\frac{2}{\pi}} \int_0^\infty e^{-\frac{1}{2}(V^2 - 2tV)} dv$$

وباكمال المربع للكمية داخل القوسين نحصل على

$$M_V(t) = \sqrt{\frac{2}{\pi}} e^{\frac{1}{2}t^2} \int_0^\infty e^{-\frac{1}{2}(V-t)^2} dv$$

الان بفرض ان $u = v - t$ اذن $-t < u < \infty$ وان $dv = du$ ، عليه فان .

$$M_V(t) = \sqrt{\frac{2}{\pi}} . e^{\frac{1}{2}t^2} \int_{-t}^\infty e^{-\frac{1}{2}u^2} du$$

$$= 2e^{\frac{1}{2}t^2} \int_{-t}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$

وكأن $u \sim N(0,1)$

$$= 2e^{\frac{1}{2}t^2} . [F(\infty) - F(-t)]$$

$$\therefore M_V(t) = 2.M_Z(t).F(t) \;;F(\infty) = 1, F(-t) = 1 - F(t)$$

لاحظ ان

$$M_V(0) = 2(1)F(0) = 2(1)\left(\frac{1}{2}\right) = 1$$

وان

$$M'_V(t) = 2[M_Z(t).F'(t) + F(t).M'_Z(t)]$$

$$= 2[M_Z(t).f(t) + F(t).M'_Z(t)]$$

فاذن

$$M'_V(0) = 2[M_Z(0).f(0) + F(0).M'_Z(0)]$$

وحيث ان $Z \sim N(0,1)$ فاذن $M'_Z(0) = EZ = 0$ وان $M_Z(0) = 1$

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$$ فاذن $$f(0) = \frac{1}{\sqrt{2\pi}}$$ وهذا يعني ان

$$M_V'(0) = 2 . \frac{1}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}} = \mu_V$$

ويطلب من القاريء اثبات ان $M_V''(0) = EV^2 = 1$

٤ ـ الدالة التوزيعية

يمكن اشتقاق صيغة للدالة التوزيعية لهذا التوزيع بالعلاقة مع الدالة التوزيعية لتوزيع $N(0,1)$ وعلى النحو التالي :

$$G(V_0) = P_r(V \leq V_0) = \sqrt{\frac{2}{\pi}} \int_0^{V_0} e^{-\frac{1}{2}V^2} dV$$

$$= 2\int_0^{V_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}V^2} dV$$

$$= 2[F(V_0) - F(0)]$$

حيث $F(.)$ هي الدالة التوزيعية لتوزيع $N(0,1)$. وحيث ان $F(0) = \frac{1}{2}$ فان

$$G(V_0) = 2\left(F(V_0) - \frac{1}{2}\right) = 2F(V_0) - 1 , 0 < V_0 < \infty$$

واذا تم صياغة $F(V_0)$ وفق الصيغة المقترحة من قبل موران الموضحة في الفقرة (٦ ـ ٢ ـ ٩) فانه يمكن البيان وبسهولة ان

$$G(V_0) = \frac{2}{\pi}\left[\frac{V_0}{3\sqrt{2}} + \sum_{n=1}^{\infty} \frac{e^{-\frac{n^2}{9}}}{n} . \sin\frac{nV_0\sqrt{2}}{3}\right]$$

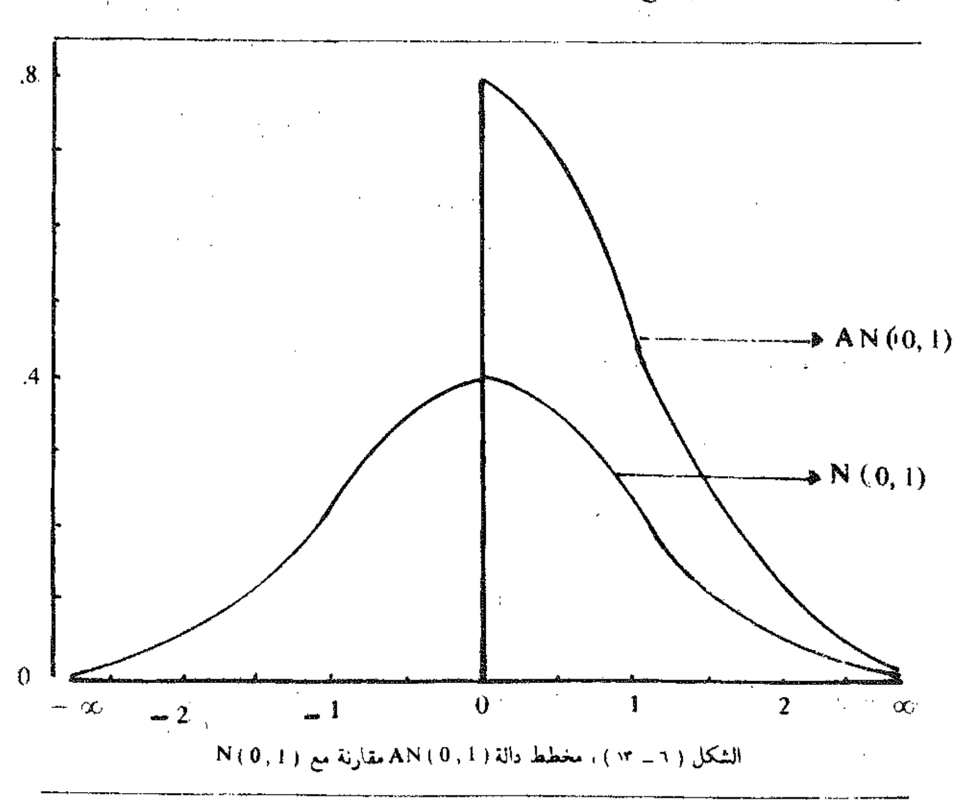
وعلى اية حال و بهدف حساب $G(1)$ مثلاً فان ذلك يتم وفق مايلي

$$G(1) = 2F(1) - 1 = 2(0.8413) - 1 = 0.6826$$

، وان

$$G(2) = 2F(\bar{2}) - 1 = 2(0.9772) - 1 = 0.9544$$

والشكل (٦ ـ ١٣) يوضح مخطط دالة $AN(0,1)$ مقارنة مع $N(0,1)$



الشكل (٦ ـ ١٣) ، مخطط دالة $AN(0,1)$ مقارنة مع $N(0,1)$

٦ ـ ٢ ـ ١٢ : أمثلة

مثال (١) : اذا كان $X \sim N(10, 16)$ فان :

1 - $f(x) = \dfrac{1}{4\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-10}{4}\right)^2}, -\infty < x < \infty$

2 - $\mu_x = 10 , \sigma_x^2 = 16$

3 - $M_X(t) = e^{10t + 8t^2}$

4 - $\text{Max}\, f(x) = \dfrac{1}{4\sqrt{2\pi}}$

5 - inflexion points are :

$\mu - \sigma = 6 , \mu + \sigma = 14$

6 - $P_r(\mu - 2\sigma < X < \mu + 2\sigma) = P_r(2 < X < 18) = 0{\cdot}9544$

7 - $P_r(X < 15) = P_r(Z < 1{\cdot}25) = 0{\cdot}8944$

8 - $P_r(X < 8) = P_r(Z < -0{\cdot}5) = 0{\cdot}3085$

والشكل (٦ ـ ١٤) يوضح مخطط دالة $N(10, 16)$

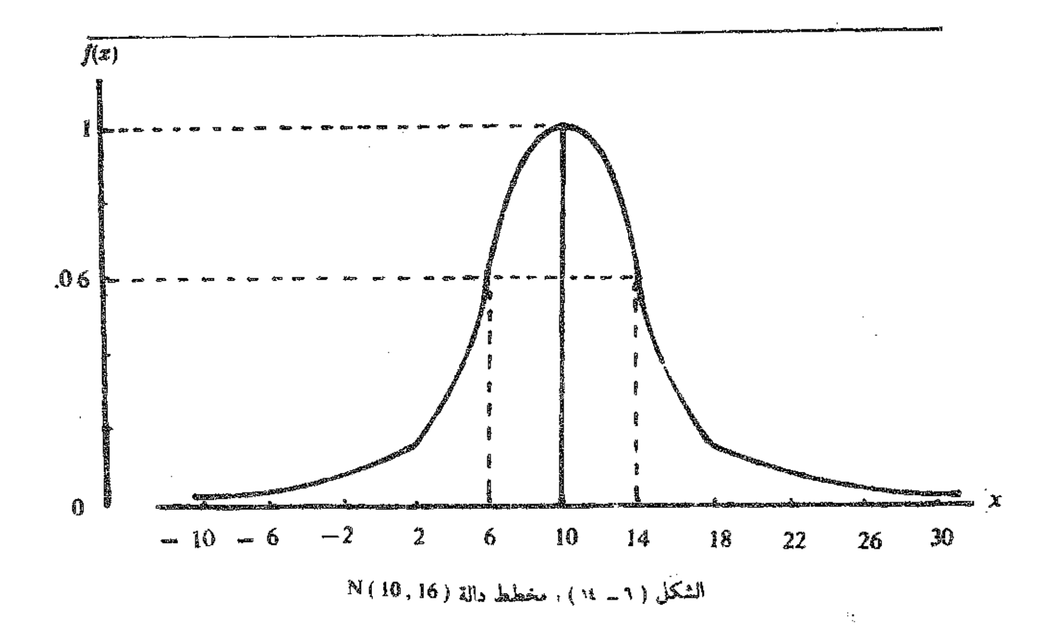
مثال (٢) : افرض ان $f(x) = ce^{-2x^2 + 8x} , -\infty < x < \infty$ جد قيمة c

الحل

$f(x) = ce^{-2(x^2 - 4x)}$

وباكمال المربع داخل القوس نحصل على :

$f(x) = ce^8 . e^{-2(x-2)^2} = ce^8 . e^{-\frac{1}{2}\left(\frac{x-2}{0{\cdot}5}\right)^2}$



الشكل (٦ ـ ١٤) ، مخطط دالة N (10 , 16)

وبالتشبيه مع دالة التوزيع الطبيعي نلاحظ ان

$$\frac{1}{\sqrt{2\pi}\ \sigma} = ce^{8}\ ,\ \mu = 2\ ,\ \sigma^{2} = \frac{1}{4}$$

$$\therefore c = \frac{1}{\frac{1}{2}\sqrt{2\pi}\ .\ e^{8}}\ ,\ X \sim N\left(2\ ,\ \frac{1}{4}\right)$$

مثال (٣)؛ لجدول التوزيع التكراري التالي يطلب توفيق توزيع طبيعي.

الفئات	0 –	10 –	20 –	30 –	40 –	50 –	60 –	70 –	80 –	90 – 100
التكرار	2	4	8	16	25	60	42	35	18	10

الحل :

ان الوسط الحسابي والتباين لهذا التوزيع هما :

$$\bar{x} = 59.5 , S^2 = 333.841 , S = 18.271$$

ونفرض ان $X \sim N(59.5 , 333.841)$. نجد الحدود الدنيا للفئات X_i ثم نحسب الدرجات المعيارية المقابلة لهذه الحدود . اي $Z_i = \frac{X_i - \bar{X}}{S}$. بعد ذلك وباستخدام جداول التوزيع الطبيعي نجد قيمة الاحتمال المتراكم لغاية Z_i . اي $F(Z_i)$ ومن ثم يتم حساب الفرق مابين الاحتمال المتراكم اللاحق والسابق له . اي $\Delta_i = F(Z_{i+1}) - F(Z_i)$ ليشكل ذلك احتمال الفئة i . بعد ذلك يتم ضرب Δ_i بمجموع التكرارات كي نحصل على التكرارات المتوقعة المقابلة لفئات التوزيع . والجدول التالي يوضح هذه العمليات :

الفئات	التكرار f_i	الحدود الدنيا للفئات X_i	الدرجات المعيارية Z_i	$F(Z_i)$	Δ_i	التكرار المتوقع $E_i = \Delta_i \Sigma f_i$
اقل من 0	–	$-\infty$	$-\infty$	0	0.0005	$0.11 \simeq 0$
0 –	2	0	– 3.26	0.0005	0.0029	$0.64 \simeq 1$
10 –	4	10	– 2.71	0.0034	0.0120	$2.64 \simeq 3$
20 –	8	20	– 2.16	0.0154	0.0383	$8.43 \simeq 8$
30 –	16	30	– 1.61	0.0537	0.0886	$19.49 \simeq 20$
40 –	25	40	– 1.07	0.1423	0.1592	$35.02 \simeq 35$
50 –	60	50	– 0.52	0.3015	0.2105	$46.31 \simeq 46$
60 –	42	60	0.03	0.5120	0.2037	$44.81 \simeq 45$
70 –	35	70	0.57	0.7157	0.1529	$33.64 \simeq 34$
80 –	18	80	1.12	0.8686	0.0839	$18.46 \simeq 18$
90 – 100	10	90	1.67	0.9525	0.0475	$10.45 \simeq 10$
اكثر من 100	–	100 فاكثر	∞	1	–	–
المجموع	220	–	–	–	–	220

مثال (٤) : اذاعلمت ان $X_1 \sim N(4,6)$, $X_2 \sim N(8,9)$, $X_3 \sim N(10,10)$
وان هذه المتغيرات مستقلة تصادفياً . عندئذٍ :

$$1-\ (X_1 + X_2 + X_3) \sim N(22, 25)$$
$$2-\ (X_1 - X_2 - X_3) \sim N(-14, 25)$$
$$3-\ (2X_1 - X_2 + 3X_3) \sim N(30, 123)$$

مثال (٥) : لتكن $X_1, X_2, \ldots, X_n$ قياسات عينة عشوائية مختارة من $N(\mu, \sigma^2)$ وليكن $\bar{X}$ يمثل الوسط الحسابي لقياسات هذه العينة . برهن ان $Z = \sqrt{n}\,(\bar{X} - \mu)/\sigma$ يتوزع كتوزيع $N(0,1)$.

البرهان :

حيث ان القياسات $X_1, X_2, \ldots, X_n$ تمثل عينة عشوائية عن متغير عشوائي X يتوزع وفق دالة توزيع $N(\mu, \sigma^2)$. فذلك يعني ان $\bar{X} \sim N(\mu, \sigma^2/n)$ عليه فان

$$Z = \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}} \sim N(0,1)$$

او ان

$$Z = \frac{\sqrt{n}\,(\bar{X} - \mu)}{\sigma} \sim N(0,1)$$

مثال (٦) : في نموذج انحدار خطي بسيط مثل $Y = \alpha + \beta X + u$ حيث ان Y يمثل المتغير التابع ، X يمثل المتغير المستقل وان $u \sim N(0, \sigma^2)$ جد التوزيع الاحتمالي للمتغير Y بفرض ثبات X علماً ان α, β عددان حقيقيان .

الحل : واضح ان المقدار $\alpha + \beta X$ كمية ثابتة ولتكن مثلاً C .
فاذن $Y = C + u$ وهذا يمثل تحويل خطي بدلالة المتغير u . وان $u \sim N(0, \sigma^2)$. عليه نستنتج ان

$$Y \sim N(EY, V(Y))$$

اي ان

$$Y \sim N(C, \sigma^2)$$

او ان

$$Y \sim N(\alpha + \beta X, \sigma^2)$$

تمارين عن التوزيع الطبيعي

٦ ـ ٥ : افرض ان $X_1 \sim N(20, 25), X_2 \sim N(20, 36), X_3 \sim N(20, 49)$
وان هذه المتغيرات مستقلة تصادفياً . يطلب اجراء مايلي :
أ ـ رسم مخطط دالة كل متغير من هذه المتغيرات على نفس الورقة مع اجراء مقارنة بينهما .
ب ـ جد التوزيع الاحتمالي للمتغير $Y = 3X_1 - 4X_2 + 2X_3$
جـ ـ جد الدالة المولدة لعزوم $Y = 2X_1 - X_2 + 4X_3$.
د ـ جد $P_r(X_1 - 2X_2 < 5), P_r(X_1 + X_3 > 2X_1), P_r(X_1 + X_2 < 2X_1 + X_3)$

٦ ـ ٦ : اذا علمت ان $X \sim N(12, 16)$. جد مايلي :
أ ـ $P_r(X > 20), P_r(0 < X < 12)$.
ب ـ جد قيمة a, b بحيث ان $P_r(a < X < b) = 0.95$ وان $P_r(X < a) = 0.05$

٦ ـ ٧ : اذا كان $Z_1 \sim N(0, 1)$ مستقل عن $Z_2 \sim N(0, 1)$. برهن ان $V_1 = Z_1 + Z_2$ مستقل عن $V_2 = Z_1 - Z_2$. جد توزيع $V = aZ_1 + bZ_2 + C$ حيث a, b, c ثوابت .

٦ ـ ٨ : اذا كان $X \sim N(\mu, \sigma^2)$. بين ان :
أ ـ الانحراف المتوسط في هذا التوزيع هو $\sigma \sqrt{\frac{2}{\pi}}$.
ب ـ ان الربيع الاول هو $Q_1 \simeq \mu - 0.67\sigma$ والربيع الثالث هو $Q_3 \simeq \mu + 0.67\sigma$.

٦ ـ ٩ : اذا كان $Z \sim N(0, 1)$. برهن ان $F(-z) = 1 - F(z)$, $z > 0$.

٦ ـ ١٠ : اذا كان $X \sim N(\mu, \sigma^2)$. جد الوسط والتباين للمتغير $Y = e^X$.

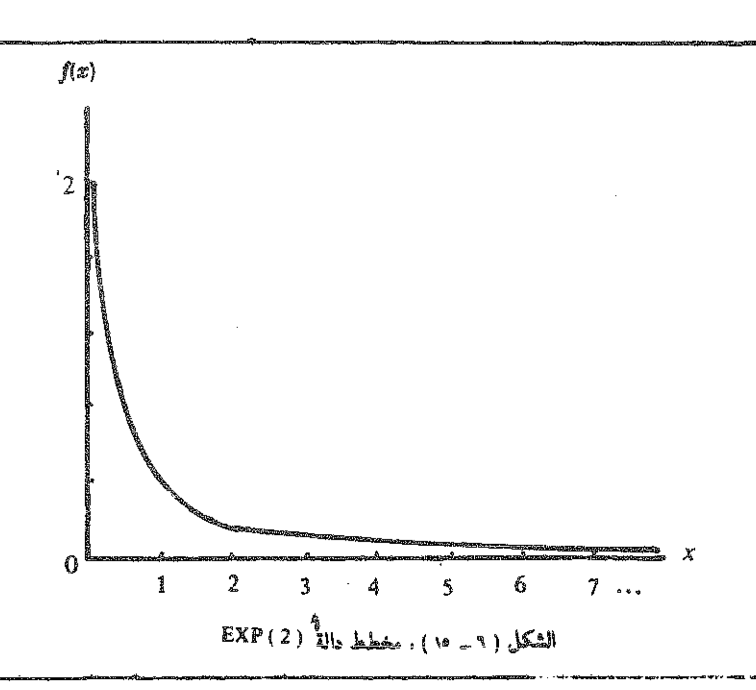
٩ ـ ٣ : التوزيع الاسي Exponential distribution

يقال للمتغير العشوائي X بانه ذو توزيع اسي اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

$$f(x;\theta) = \theta e^{-\theta x} \qquad ; x \geq 0$$
$$= 0 \qquad \text{other wise}$$

حيث θ تمثل معلمة التوزيع وان $\theta > 0$. وبالرموز فان $X \sim EXP(\theta)$

والشكل (٦ ـ ١٥) يوضح مخطط دالة EXP (2) .



الشكل (٩ ـ ١٥) : مخطط دالة EXP (2)

ويمكن اثبات ان المساحة تحت منحنى دالة هذا التوزيع مساوية للواحد دلالة على كون f (x) هي دالة كثافة احتمالية وكما يلي :

$$\int_0^{\infty} f(x;\theta)\,dx = \int_0^{\infty} \theta e^{-\theta x}\,dx = -\left[e^{-\theta x}\right]_0^{\infty} = 1$$

٦ ـ ٣ ـ ١ : الدالة التوزيعية :

الدالة التوزيعية في التوزيع الاسي هي :

$$F(x) = P_r(X \le x) = \int_0^x \theta e^{-\theta u}\, du = -[e^{-\theta u}]_0^x$$

$$= 1 - e^{-\theta x} \quad ; \quad 0 < x < \infty$$

ونظراً لسهولة حساب $F(x)$ وفق الصيغة اعلاه فانه ليس من الضروري بناء جداول خاصة بهذا التوزيع . فمثلاً لو كانت $\theta = 1$ فان :

$$F(1) = 0{\cdot}6321206 ,\ F(2) = 0{\cdot}8646648 ,\ F(3) = 0{\cdot}9502130$$

٦ ـ ٣ ـ ٢ : الدالة المولدة للعزوم حول نقطة الاصل .

ان الدالة المولدة لعزوم التوزيع الاسي حول نقطة الاصل هي :

$$M_X(t) = Ee^{tX} = \theta \int_0^\infty e^{tx} . e^{-\theta x}\, dx$$

$$= \theta \int_0^\infty e^{-x(\theta - t)}\, dx$$

$$= \frac{\theta}{\theta - t} , t < \theta$$

ومن خلال هذه الدالة يمكن استنتاج عزوم التوزيع حول نقطة الاصل وكما يلي :

$$M_X'(t) = \frac{\theta}{(\theta - t)^2} \qquad \therefore M_X'(0) = \mu_x = \frac{1}{\theta}$$

$$M_X''(t) = \frac{2\theta}{(\theta - t)^3} \qquad \therefore M_X''(0) = EX^2 = \frac{2}{\theta^2}$$

فاذن

$$\sigma_X^2 = \frac{1}{\theta^2} - \frac{1}{\theta^2} = \frac{1}{\theta^2}$$

كذلك فان

$$M_X'''(t) = \frac{6\theta}{(\theta - t)^4} \qquad \therefore M_X'''(0) = EX^3 = \frac{6}{\theta^3}$$

وبشكل عام فان

$$M_X^{(r)}(t) = \frac{r!\theta}{(\theta - t)^{r+1}} \rightarrow M_X^{(r)}(0) = \frac{r!}{\theta^r}, r = 1, 2, \ldots$$

٦ ـ ٣ ـ ٣ : الوسيط في التوزيع الاسي .

بشكل عام فان الوسيط يمثل قيمة من قيم المتغير العشوائي X التي تحقق $F(x) = \frac{1}{2}$. وهذا يعني ان :

$$1 - e^{-\theta x} = \frac{1}{2} \rightarrow e^{-\theta x} = \frac{1}{2}$$

او ان

$$-\theta x = \ln\left(\frac{1}{2}\right) = -\ln 2$$

اذن

$$x = \frac{\ln 2}{\theta}$$

ويلاحظ مما تقدم ان قيمة الوسيط تزداد اقتراباً من الصفر كلما ابتعدت θ عن الصفر نحو الجانب الموجب .

٦ ـ ٣ ـ ٤ : امثلة

مثال (١) : اذا كان $X \sim EXP\ 4$ فان :

$1 - \mu_x = \frac{1}{4}, \sigma_x^2 = \frac{1}{16}$

$2 - EX^r = \frac{r!}{4^r}, r = 1, 2, \ldots$

$3 - x = 0{\cdot}1732867$ is the median of X .

$4 - F(1) = 0{\cdot}9816844$

مثال (٢) : لجدول التوزيع التكراري التالي يطلب توفيق توزيع اسي .

الفئات :	0	10 –	20 –	30 –	40 –	50 –	60 – 70
التكرار :	100	60	20	10	7	3	0

الحل :

نجد الوسط الحسابي لهذا التوزيع . اي .

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = 13{\cdot}65$$

ونجعل الوسط الحسابي لهذا التوزيع التكراري مساوياً لمتوسط التوزيع الاسي . اي

$$\mu_x = \frac{1}{\theta} = 13{\cdot}65 \qquad \therefore \theta = 0{\cdot}07326$$

ونفرض ان $X \sim EXP(0{\cdot}07326)$. ونعمل الجدول التالي الذي يمثل خلاصة العمليات الحسابية التي تقودنا الى التكرارات المتوقعة .

الفئات	التكرار f_i	الحدود العليا للفئات X_i	$F(x_i)$	احتمال الفئة $\Delta F(x_i)$	التكرار المتوقع $\Delta F(x_i)\Sigma f_i$
0 –	100	10	0·51934	0·51934	104
10 –	60	20	0·76897	0·24963	50
20 –	20	30	0·88895	0·11998	24
30 –	10	40	0·94662	0·05767	12
40 –	7	50	0·97434	0·02772	6
50 –	3	60	0·98767	0·01333	3
60 – 70	0	70	0·99407	0·00640	1
المجموع	200	–	–	0·99407 ≃ 1	200

مثال (٣) : اذا علمت ان X يتوزع كتوزيع منتظم مستمر على الفترة $(0,1)$ وان $Y = -\frac{1}{\theta}\ln X$ ، برهن ان $Y \sim EXP(\theta)$.

الحل :

لتكن $G(y)$ تمثل الدالة التوزيعية للمتغير Y ، وهذا يعني ان :

$$G(y) = P_r(Y \leq y) = P_r\left(-\frac{1}{\theta}\ln X \leq y\right)$$
$$= P_r(\ln X \geq -\theta y) = P_r(X \geq e^{-\theta y})$$
$$= \int_{e^{-\theta y}}^{1} dx = 1 - e^{-\theta y}$$

وبتفاضل الطرفين نسبة للمتغير Y نحصل على :

$$g(y) = G'(y) = \theta e^{-\theta y}$$

وعلى اساس التحويل $Y = -\frac{1}{\theta}\ln X$ يلاحظ مايلي :

عندنا $x \to 0$ فان $y \to \infty$ ، وعندما $x \to 1$ فان $y \to 0$ وهذا يعني ان $0 < y < \infty$. ومما تقدم نستنتج ان $Y \sim EXP(\theta)$.

تمارين عن التوزيع الاسي

٦ ــ ١١ :افرض ان $X \sim EXP(5)$ يطلب اجراء مايلي :

أ ــ رسم مخطط دالة هذا التوزيع .

ب ــ جد الوسط والتباين للمتغير X .

جـ ــ جد الوسيط في هذا التوزيع .

د ــ ماهو نوع التواء هذا التوزيع استناداً لمعطيات الفرعين ب . جـ .

هـ ــ جد الدالة المولدة لعزوم $Y = 2X + 3$ حول نقطة الاصل .

٦ ــ ١٢ : اذا كان $X \sim EXP(\theta)$. برهن ان $P_r(X \leq x \mid x \geq a) = F(x)$
حيث $Y = X - a, a \geq 0$.

٦ ــ ١٣ : اذا كان $X \sim EXP(\theta)$. برهن ان $M_X(t) = \sum_{k=0}^{\infty} \left(\frac{t}{\theta} \right)^k$

٦ ــ ١٤ : اشتق صيغة للدالة المولدة للعزوم المركزية لتوزيع $EXP(\theta)$.

٦ ــ ١٥ : اذا كان $X \sim EXP(\theta)$ جد الوسط والتباين للمتغير $Y = X^n$.

٦ ـ ٤ : توزيع كاما Gamma distribution

ان توزيع كاما في الحقيقة مشتق من دالة كاما Gamma function او ماتسمى في بعض الاحيان بتكامل كاما الذي يرد ذكره في الكثير من كتب الرياضيات المتقدمة . ويمكن عد هذا التوزيع كواحد من التوزيعات المهمة في دراسة المشكلات التي يكون الزمن احد عواملها كتلك الدراسات الخاصة بطول مدة اشتغال معدات مصنع معين . دراسة العطلات والتوقفات لمكائن مصنع معين . كذلك يعد من التوزيعات المهمة التي تدخل في دراسة موضوع المعولية reliability . ويعرف تكامل كاما رياضا بالشكل التالي :

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} . e^{-y} dy$$

حيث $\Gamma(\alpha)$ تمثل قيمة تكامل كاما عند قيمة معينة الى α وان هذا التكامل متقارب لجميع قيم $\alpha > 0$ ومتباعدلقيم $\alpha \leq 0$. فمثلا عندما $\alpha = 1$ فان

$$\Gamma(1) = \int_0^{\infty} e^{-y} dy = 1$$

لاحظ هنا ان التكامل متقارب . واذا كانت $\alpha = 0$ فان

$$\Gamma(0) = \int_0^{\infty} y^{-1} e^{-y} dy$$

وباستخدام التكامل بالتجزئة نلاحظ ان

$$\Gamma(0) = \lim_{y \to \infty} e^{-y} \ln y - \lim_{y \to 0} e^{-y} \ln y$$

وهذا يعني ان $\Gamma(0)$ شكل غير محدد اي ان التكامل متباعد . كذلك فان $\Gamma(\alpha), \alpha > 0$ عدد موجب . وبقسمة طرفي دالة كاما على $\Gamma(\alpha)$ نحصل على :

$$1 = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} y^{\alpha-1} e^{-y} dy$$

وهذا يعني ان Y متغير عشوائي بدالة كثافة احتمالية

$$f(y;\alpha) = \frac{1}{\Gamma(\alpha)} y^{\alpha-1}e^{-y} ; y > 0$$

$$= 0 \qquad \text{otherwise}$$

وهذا يسمى الشكل الاول لدالة توزيع كاما بالمعلمة $\alpha > 0$. وبالرموز فان $Y \sim G(\alpha)$. ويتضح من ذلك ان التوزيع الاسي بالمعلمة $\theta = 1$ هو حالة خاصة من توزيع كاما عندما $\alpha = 1$. وهنالك شكل آخر لدالة توزيع كاما مشتق من الشكل الاول وهو الآتي :

بفرض ان $Y = \frac{X}{\beta}$ حيث $\beta > 0$ عندئذ

$$\Gamma(\alpha) = \int^{\infty} \left(\frac{x}{\beta}\right)^{\alpha-1} . e^{-\frac{x}{\beta}} . \frac{1}{\beta} dx$$

$$= \frac{1}{\beta^{\alpha}} \int_{0}^{\infty} x^{\alpha-1} . e^{-\frac{x}{\beta}} dx$$

وبقسمة الطرفين على $\Gamma(\alpha)$ نحصل على :

$$1 = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int x^{\alpha-1} e^{-\frac{x}{\beta}} ; x > 0$$

وهذا يعني ان X متغير عشوائي بدالة كثافة احتمالية :

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1}e^{-\frac{x}{\beta}} , x > 0$$

$$= 0 \qquad \text{other wise}$$

وهذا هو الشكل الثاني لدالة توزيع كاما بالمعلمتين $\alpha , \beta > 0$ وبالرموز فان $X \sim G(\alpha, \beta)$. ويتضح من هذا الشكل ان التوزيع الاسي هو حالة خاصة من توزيع كاما عندما $\alpha = 1, \theta = \frac{1}{\beta}$. ويمكن اثبات ان قيمة تكامل كاما هي $\Gamma(\alpha) = (\alpha - 1)!$ وكما يلي :

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha - 1} e^{-y} dy$$

وباستخدام التكامل بالتجزئة من خلال الفرض ان $dv = e^{-y}dy , u = y^{\alpha - 1}$ عندئذٍ $V = - e^{-y}, du = (\alpha - 1) y^{\alpha - 2} dy$. فاذن

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha - 1} e^{-y} dy = [- y^{\alpha - 1} e^{-y}]_0^\infty + (\alpha - 1) \int_0^\infty y^{\alpha - 2} e^{-y} dy$$

$$= (\alpha - 1) \int_0^\infty y^{\alpha - 2} e^{-y} dy$$

$$= (\alpha - 1) \Gamma(\alpha - 1)$$

وبنفس الاجراء السابق يمكن ملاحظة ان

$$\Gamma(\alpha - 1) = (\alpha - 2) \Gamma(\alpha - 2)$$

اي ان

$$\Gamma(\alpha) = (\alpha - 1)(\alpha - 2) \Gamma(\alpha - 2)$$

واذا استمر الحال باجراء التكامل بالتجزئة فان

$$\begin{aligned} \Gamma(\alpha) &= (\alpha - 1)(\alpha - 2)(\alpha - 3) \ldots 3 \cdot 2 \cdot 1 \cdot 0! \\ &= (\alpha - 1)! \end{aligned}$$

كذلك يمكن اثبات ان $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ على النحو الاتي :

ان

$$\Gamma\left(\frac{1}{2}\right)=\int_0^{\infty} y^{-\frac{1}{2}} e^{-y}dy$$

وبفرض ان $y = \frac{1}{2}x^2$ فان $dy = xdx$. عليه فان

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty}\left(\frac{1}{2}x^2\right)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2}.x\,dx$$

$$= \sqrt{2}\int_0^{\infty} e^{-\frac{1}{2}x^2}dx$$

لكن :

$$\int_0^{\infty} e^{-\frac{1}{2}x^2}dx = \frac{1}{2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}dx$$

$$= \frac{1}{2}\sqrt{2\pi} = \sqrt{\frac{\pi}{2}}$$

فاذن

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{2}\cdot\sqrt{\frac{\pi}{2}} = \sqrt{\pi}$$

واستنادا لهذه العلاقة فانه لاي عدد صحيح موجب مثل r فان

$$\Gamma\left(\frac{2r+1}{2}\right) = \frac{1.\,3.\,5.\,7.\,....\,(2r-1)}{2^r}\sqrt{\pi}$$

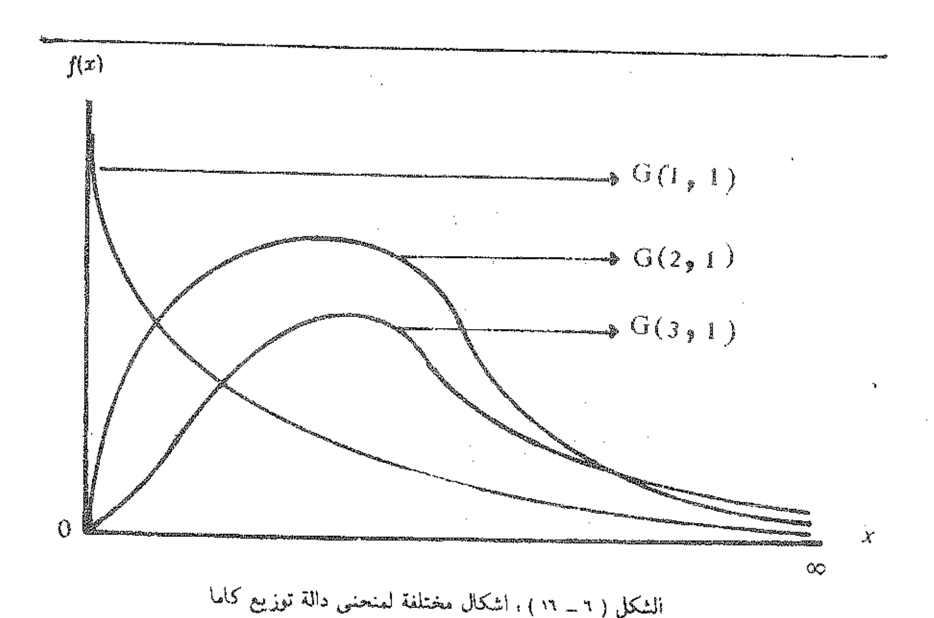
فمثلا عندما :

$$r = 0 \quad \rightarrow \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$r = 1 \quad \rightarrow \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}\sqrt{\pi}$$

$$r = 2 \quad \rightarrow \Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{3}{4}\sqrt{\pi}$$

وهكذا الحال بالنسبة لاية قيمة اخرى الى r . والشكل (٦ ـ ١٦) يوضح مخطط لثلاثة توزيعات كاما .



الشكل (٦ ـ ١٦) ، اشكال مختلفة لمنحنى دالة توزيع كاما

٦ ـ ٤ ـ ١ : الدالة التوزيعية :

تعرف الدالة التوزيعية في توزيع كاما كالاتي :

$$F(y) = P_r(Y \leq y) = \frac{1}{\Gamma(\alpha)} \int_0^y z^{\alpha-1} . e^{-z} \, dz$$

ان هذا التكامل هو تكامل كاما الناقص المنوه عنه في الفقرة (٥ ـ ٦) لدى دراستنا للدالة التوزيعية في توزيع بواسون . وهذا يعني أن

$$F(y) = 1 - P_r(Y > y) = 1 - \frac{1}{\Gamma(\alpha)} \int_y^{\infty} z^{\alpha-1} . e^{-z} dz$$

$$= 1 - \sum_{j=0}^{\alpha-1} \frac{y^j e^{-y}}{j!}$$

اي ان

$$F(y) = \sum_{j=\alpha}^{\infty} \frac{y^j e^{-y}}{j!}$$

ويلاحظ من الصيغة الاخيرة انها ممكنة التطبيق في حالة α عدد موجب صحيح . وسوف لن نتوسع في دراسة هذه الدالة هنا بسبب ضيق حيز تطبيقها من جهة اضافة الى اننا سوف نعرض وبشكل مفصل حالة خاصة لها لدى دراستنا لتوزيع مربع كاي في الفقرة (٩ ـ ١ ـ ٣) .

٦ ـ ٤ ـ ٢ : الدالة المولدة لعزوم التوزيع حول نقطة الاصل .

ان الدالة المولدة لعزوم توزيع كاما حول نقطة الاصل هي

$$M_X(t) = (1 - \beta t)^{-\alpha} .$$

البرهان : ان

$$M_X(t) = Ee^{tX} = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} e^{tx} . x^{\alpha-1} . e^{-\frac{x}{\beta}} dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} x^{\alpha-1} . e^{-x\left(\frac{1-\beta t}{\beta}\right)} dx$$

الان بفرض ان $\beta^* = \frac{\beta}{1-\beta t}$ فان :

$$M_X(t) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^{\alpha-1}.e^{-\frac{x}{\beta^*}}\,dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha}.\Gamma(\alpha)\beta^{*\alpha} = \frac{1}{\beta^\alpha}.\left(\frac{\beta}{1-\beta t}\right)^\alpha$$

$$= (1-\beta t)^{-\alpha}, t < \beta^{-1}$$

مع ملاحظة انه اذا كانت $\beta = 1$ فان $M_X(t) = (1-t)^{-\alpha}$ التي تمثل الدالة المولده لعزوم توزيع كاما بشكله الاول .

وعن طريق هذه الدالة يمكن استنتاج عزوم التوزيع وعلى النحو التالي :

$$M'_X(t) = \alpha\beta(1-\beta t)^{-\alpha-1} \rightarrow M'_X(0) = \alpha\beta = \mu_x$$

وان

$$M''_X(t) = \alpha(\alpha+1)\beta^2(1-\beta t)^{-\alpha-2} \rightarrow M''_X(0) = \alpha(\alpha+1)\beta^2 = EX^2$$

فاذن

$$\sigma_x^2 = \alpha(\alpha+1)\beta^2 - \alpha^2\beta^2 = \alpha\beta^2$$

واذا كانت $\beta = 1$ فان $\mu_x = \sigma_x^2 = \alpha$.

وبشكل عام فان العزم ذو المرتبة r حول نقطة الاصل هو :

$$EX^r = \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^r.x^{\alpha-1}.e^{-\frac{x}{\beta}}\,dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^{(r+\alpha)-1}.e^{-\frac{x}{\beta}}\,dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha}.\Gamma(r+\alpha).\beta^{r+\alpha}$$

$$= \beta^r.\frac{\Gamma(r+\alpha)}{\Gamma(\alpha)}, r = 1, 2, \ldots$$

وهذا يعني ان

$$M_X^{(r)}(0) = EX^r = \beta^r . \frac{\Gamma(r+\alpha)}{\Gamma(\alpha)}$$

٦ ـ ٤ ـ ٣ : خاصية الجمع في توزيع كاما Additive property

اذا كانت $X_1, X_2, \ldots, X_k$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim G(\alpha_i, \beta)$ عندئذٍ فان $Y = \sum_{i=1}^{k} X_i \sim G\left(\sum_{i=1}^{k} \alpha_i, \beta\right)$.

البرهان : افرض ان المتغير Y يمتلك دالة مولدة للعزوم حول نقطة الاصل . وهذا يعني ان :

$$M_Y(t) = Ee^{tY} = Ee^{t\sum_{i=1}^{k} X_i}$$

$$= E\prod_{i=1}^{k} e^{tX_i} = \prod_{i=1}^{k} Ee^{tX_i} = \prod_{i=1}^{k} M_{X_i}(t)$$

وحيث ان $X_i \sim G(\alpha_i, \beta)$ فذلك يعني ان $M_{X_i}(t) = (1-\beta t)^{-\alpha_i}$. عليه فان

$$M_Y(t) \quad \prod_{i=1}^{k} (1-\beta t)^{-\alpha_i} = (1-\beta t)^{-\sum_{i=1}^{k} \alpha_i}$$

والصيغة الاخيرة تمثل الدالة المولدة لعزوم متغير عشوائي يتوزع وفق دالة توزيع كاما بالمعلمتين $\sum_{i=1}^{k} \alpha_i$, β . فاذن $\sum_{i=1}^{k} X_i \sim G\left(\sum_{i=1}^{k} \alpha_i, \beta\right)$.

٦ ـ ٤ ـ ٤ : خاصية التقارب من التوزيع الطبيعي :

افرض ان $X \sim G(\alpha)$ وان $Z = \dfrac{X-\alpha}{\sqrt{\alpha}}$ عندئذٍ فان $Z \sim N(0,1)$ عندما $\alpha \to \infty$.

البرهان : حيث ان $X \sim G(\alpha)$ فان $Z = \dfrac{X-\alpha}{\sqrt{\alpha}}$ تمثل الدرجة المعيارية في هـذا التوزيـع وان $\sigma_Z^2 = 1, \mu_Z = 0$. لتكن $M_Z(t)$ تمثل الدالة المولدة لعزوم Z . فاذن

$$M_Z(t) = Ee^{tZ} = e^{-t\sqrt{\alpha}} \cdot M_X\left(\frac{t}{\sqrt{\alpha}}\right)$$

$$= e^{-t\sqrt{\alpha}} \cdot \left(1 - \frac{t}{\sqrt{\alpha}}\right)^{-\alpha}$$

وان

$$K_Z(t) = \ln M_Z(t) = -t\sqrt{\alpha} - \alpha \ln\left(1 - \frac{t}{\sqrt{\alpha}}\right)$$

لكن وحسب متسلسلة تايلر فان

$$\ln(1+K) = K - \frac{K^2}{2} + \frac{K^3}{3} - \ldots, |K| < 1$$

واذا فرضنا ان $K = -\dfrac{t}{\sqrt{\alpha}}$ فان $|K| < 1$ عندما تقترب α من اللانهاية فاذن :

$$K_Z(t) = -t\sqrt{\alpha} + \alpha\left(\frac{t}{\sqrt{\alpha}} + \frac{t^2}{2\alpha} + \frac{t^3}{3\alpha\sqrt{\alpha}} + \ldots\right)$$

$$= \frac{t^2}{2} + \frac{t^3}{3\sqrt{\alpha}} + 0\left(\frac{1}{\alpha}\right)$$

حيث $0\left(\frac{1}{\alpha}\right)$ تمثل حدود لاحقة تتضمن α في مقاماتها . عليه فان

$$\lim_{\alpha \to \infty} K_Z(t) = \frac{t^2}{2}$$

وهذا يعني ان :

$$\lim_{\alpha \to \infty} M_Z(t) = e^{\frac{1}{2}t^2}$$

والصيغة الاخيره تمثل الدالة المولدة لعزوم التوزيع الطبيعي المعياري . وبذلك نستنتج ان $Z \sim N(0,1)$ عندما $\alpha \to \infty$ وان

$$\lim_{\alpha \to \infty} G(\alpha) \to N(\alpha, \alpha)$$

كذلك يمكن التوصل . وبنفس الاسلوب اعلاه . الى ان

$$\lim_{\alpha \to \infty} G(\alpha, \beta) \to N(\alpha\beta, \alpha\beta^2)$$

ونترك برهنة ذلك للقاريء . ان لهذه الخاصية اهمية كبيرة اثناء التطبيق . فهي تسمح لنا باستخدام جداول التوزيع الطبيعي لحساب احتمال معين في حالة ملاحظتنا ان α كبيرة .فمثلاً اذا كان $X \sim G(1000)$ وتطلب الامر حساب $P_r(X < 980)$ فان ذلك يتم على النحو التالي : حيث ان α كبيرة نسبياً فان

$$P_r(X < 980) = P_r\left(\frac{X - 1000}{\sqrt{1000}} < \frac{980 - 1000}{\sqrt{1000}}\right)$$

$$= P_r(Z < -0{\cdot}63)$$

ومن جداول التوزيع الطبيعي نلاحظ ان $F(-0{\cdot}63) = 0{\cdot}2643$ فاذن

$$P_r(X < 980) = \frac{1}{\Gamma(1000)} \int_0^{980} x^{999} \cdot e^{-x}\, dx \simeq 0{\cdot}2643$$

٦ ـ ٤ ـ ٥ : المنوال ونقاط الانقلاب في توزيع كاما

بفرض ان $X \sim G(\alpha, \beta)$ عندئذٍ يمكن اثبات ان المنوال لهذا التوزيع هو $x = \beta(\alpha - 1)$ وان هنالك نقطتا انقلاب في منحنى دالة هذا التوزيع هما $x = \beta(\alpha - 1) \pm \beta\sqrt{\alpha - 1}$.

البرهان : ان

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

$$\therefore \ln f(x) = \ln\left(\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\right) + (\alpha - 1)\ln x - \frac{x}{\beta}$$

وبايجاد المشتقة الاولى نسبة الى X نحصل على :

$$\frac{f'(x)}{f(x)} = \frac{\alpha - 1}{x} - \frac{1}{\beta} \rightarrow f'(x) = f(x)\left[\frac{\alpha - 1}{x} - \frac{1}{\beta}\right]$$

وبوضع $f'(x) = 0$ فان :

$$f(x)\left[\frac{\alpha - 1}{x} - \frac{1}{\beta}\right] = 0 \rightarrow \frac{\alpha - 1}{x} - \frac{1}{\beta} = 0, f(x) > 0$$

وهذا يعني ان $x = \beta(\alpha - 1)$ كذلك فان :

$$f''(x) = f(x)\left[-\frac{\alpha - 1}{x^2} + \left(\frac{\alpha - 1}{x} - \frac{1}{\beta}\right)\right]^2$$

وان

$$f''(x)\Big]_{x=\beta(\alpha-1)} = -\frac{1}{\beta^2(\alpha - 1)} f(x = \beta(\alpha - 1)) < 0$$

وحيث ان المشتقة الثانية سالبة عندما $x = \beta(\alpha - 1)$ فذلك يعني ان منحنى دالة هذا التوزيع له نهاية عظمى عندما $x = \beta(\alpha - 1)$، من ذلك نستنتج ان المنوال لتوزيع كاما هو $x = \beta(\alpha - 1)$

وبجعل $f''(x) = 0$ لغرض ايجاد نقاط الانقلاب . نحصل على

$$f(x).\left[-\frac{\alpha-1}{x^2}+\left(\frac{\alpha-1}{x}-\frac{1}{\beta}\right)^2\right]=0$$

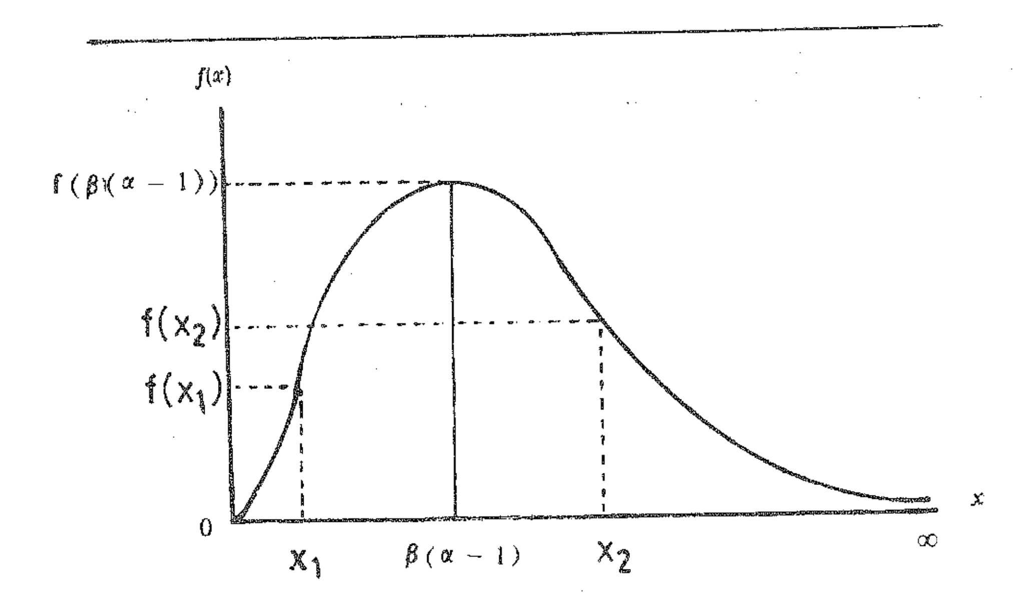
لكن $f(x) > 0$ فاذن

$$-\frac{\alpha-1}{x^2}+\left(\frac{\alpha-1}{x}-\frac{1}{\beta}\right)^2=0$$

وبحل الصيغة الاخيرة نسبة الى x نحصل على :

$$x=\beta(\alpha-1)\pm\beta\sqrt{\alpha-1} \quad ;\alpha>1$$

ونترك للقاريء اثبات انه عندما $x=\beta(\alpha-1)\pm\beta\sqrt{\alpha-1}$ فان $f'''(x) \neq 0$. وهذا يعني ان لمنحنى دالة توزيع كاما نقطتا انقلاب تقعان على بعد متساو الى يمين ويسار المنوال هما $x=\beta(\alpha-1)\pm\beta\sqrt{\alpha-1}$. والشكل (٦ ـ ١٧) يوضح موقع المنوال ونقطتي الانقلاب :



الشكل (٦ ـ ١٧) : موقع المنوال ونقطتي الانقلاب في دالة توزيع كاما .

٦ ـ ٤ ـ ٦ : امثلة

مثال (١) : اذا كان $X \sim G(2,3)$ فان :

$$1 - f(x) = \frac{1}{9} x e^{-\frac{x}{3}} \quad ; x > 0$$

$$2 - \mu_x = \alpha\beta = 6 \ , \sigma_x^2 = \alpha\beta^2 = 18$$

$$3 - M_X(t) = (1 - 3t)^{-2}, t < \frac{1}{3}$$

$$4 - EX^r = 3^r \Gamma(r+2), r = 1, 2, \ldots$$

$$5 - F(x) = \frac{1}{9} \int_0^x u e^{-\frac{u}{3}} du = 1 - \left(\frac{x}{3} + 1 \right) e^{-\frac{1}{3}x}$$

واضح ان

$$F(o) = 0, F(1) = 0{\cdot}044625, F(10) = 0{\cdot}8454132$$

مثال (٢) : اذا كان $X \sim G(\alpha, \beta)$. برهن ان الوسط التوافقي مساو للمنوال في هذا التوزيع . ثم اشتق صيغة الى EX^{-r} ، r عدد موجب صحيح .

الحل : حسب تعريف الوسط التوافقي فان :

$$\frac{1}{H} = E \frac{1}{X} = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \frac{1}{x} . x^{\alpha-1} . e^{-x/\beta} dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{(\alpha-1)-1} . e^{-x/\beta} dx$$

ويتضح ان قيمة التكامل في الصيغة الاخيرة هي $\Gamma(\alpha - 1) . \beta^{\alpha-1}$ عليه فان :

$$\frac{1}{H} = \frac{1}{\Gamma(\alpha)\beta^\alpha} . \Gamma(\alpha - 1)\beta^{\alpha-1} = \frac{1}{\beta(\alpha - 1)}$$

فاذن $H = \beta(\alpha - 1)$.

كذلك فان

$$EX^{-r} = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} x^{-r} . x^{\alpha-1} . e^{-x/\beta} dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} x^{(\alpha-r)-1} . e^{-x/\beta} dx$$

عليه فان .

$$EX^{-r} = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} . \Gamma(\alpha - r)\beta^{\alpha-r} = \frac{\Gamma(\alpha - r)}{\Gamma(\alpha) . \beta^{r}} , r < \alpha$$

مثال (٣) : افرض ان $X_1, X_2, \ldots, X_n$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim EXP(\theta)$. برهن ان $\sum_{i=1}^{n} X_i \sim G\left(n, \frac{1}{\theta}\right)$

البرهان :

افرض ان $Y = \sum_{i=1}^{n} X_i$ وان Y يمتلك دالة مولدة للعزوم حول نقطة الاصل . وحسب هذا الفرض فان :

$$M_Y(t) = Ee^{tY} = Ee^{t\sum x_i}$$

$$= E \prod_{i=1}^{n} . e^{tX_i}$$

وحيث ان المتغيرات $X_1, X_2, \ldots, X_n$ مستقلة فذلك يعني ان ،

$$M_Y(t) = \prod_{i=1}^{n} Ee^{tX_i} = \prod_{i=1}^{n} M_{X_i}(t)$$

لكن $X_i \sim EXP(\theta)$ وذلك يعني ان $M_{X_i}(t) = \frac{\theta}{\theta - t}$ ،

فاذن

$$M_Y(t) = \prod_{i=1}^{n} \frac{\theta}{\theta - t} = \left(\frac{\theta}{\theta - t}\right)^n$$

وبقسمة البسط والمقام على θ نحصل على :

$$M_Y(t) = \left(\frac{1}{1 - \frac{1}{\theta} t} \right)^n$$

ويتضح من هذه الصيغة انها تمثل الدالة المولدة لعزوم متغير مثل Y حول نقطة الاصل يتوزع كتوزيع كاما بالمعلمتين $\alpha = n$, $\beta = \frac{1}{\theta}$. فاذن

$$\sum_{i=1}^{n} X_i \sim G\left(n, \frac{1}{\theta}\right); X_i \sim EXP(\theta)$$

تمارين عن توزيع كاما

٦ ـ ١٦ : اذا كان $X_1 \sim G(2,3)$ مستقل عن $X_2 \sim G(1,3)$. جد مايلي

أ ـ دالة الكثافة الاحتمالية للمتغير $Y = X_1 + X_2$ ثم ارسم مخطط هذه الدالة .

ب ـ الوسط والتباين للمتغير Y .

جـ ـ $P_r(Y > 2)$ ، $P_r(3 < Y < 10)$.

٦ ـ ١٧ : اذا كان $\bar{X}$ يمثل الوسط الحسابي لقياسات عينة عشوائية قوامها n مسحوبة من $G(\alpha, \beta)$ برهن ان $X \sim G\left(n\alpha, \frac{\beta}{n}\right)$

٦ ـ ١٨ : برهن ان معامل الالتواء لتوزيع $G(\alpha, \beta)$ لايعتمد على β . ثم بين ان هذا التوزيع ملتو التواء موجب . واذا كانت $\alpha \to \infty$ بين ان التوزيع يقترب من حالة التماثل .

٦ ـ ١٩ : اذا كان $X \sim N(0,1)$. برهن ان $Y = \frac{1}{2} X^2 \sim G\left(\frac{1}{2}, 1\right)$

٦ ـ ٢٠ : اذا كانت $Z_1, Z_2, \ldots, Z_n$ متغيرات عشوائية مستقلة كل منها يتوزع وفق دالة توزيع $N(0,1)$. برهن ان

$$\frac{1}{2} \sum_{i=1}^{n} Z_i^2 \sim G\left(\frac{n}{2}, 1\right)$$

٦ ـ ٢١ : اذا كان $X \sim G(\alpha, 1)$ برهن ان $EX^r = \prod_{j=1}^{r} (\alpha + j - 1), r = 1, 2, \ldots$

٦ ـ ٥ : توزيع بيتا Beta distribution

ان هذا التوزيع مشتق من دالة بيتا Beta function او ماتسمى في بعض الاحيان تكامل بيتا . ويعد واحداً من التوزيعات ذات اهمية تطبيقية في حقل الرقابة على جودة الانتاج من خلال تكوين مايسمى « جداول عينات القبول Acceptance sampling tables » التي تستخدم في اتخاذ القرار بشأن قبول وجبات الانتاج استناداً الى نسب الوحدات المعيبة في العينة.ان تكامل بيتا معطى بالصيغة التالية :

$$B(\alpha,\beta)=\int_0^1 x^{\alpha-1}.(1-x)^{\beta-1}\,dx$$

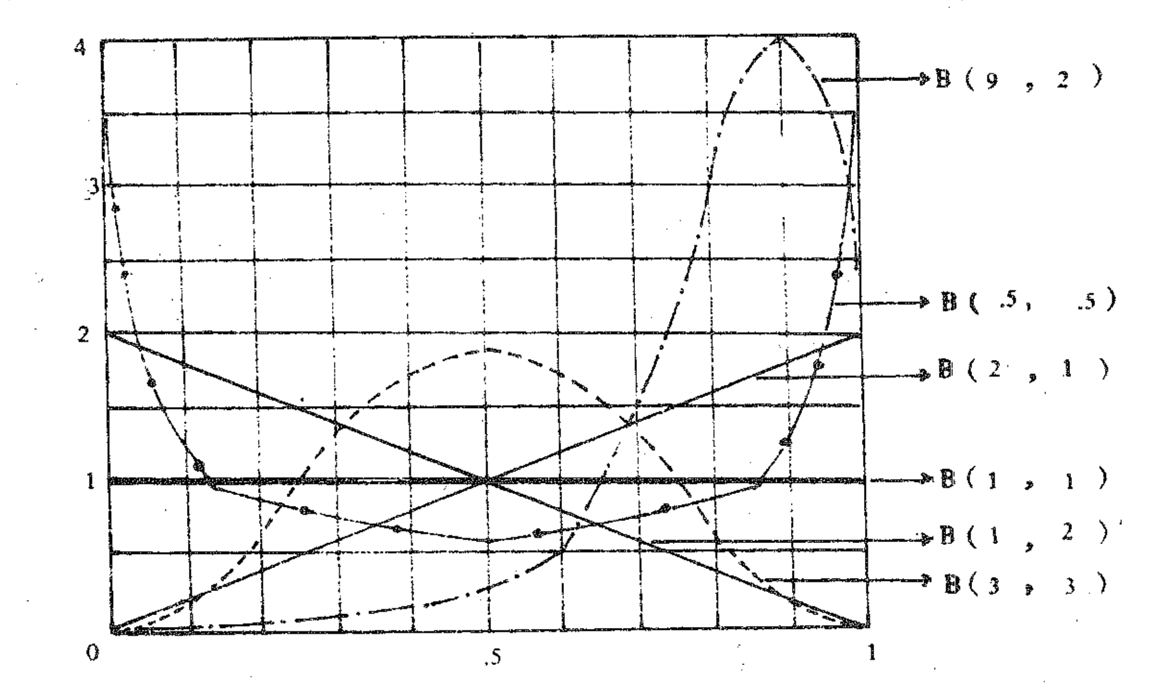
حيث $B(\alpha,\beta)=\dfrac{\Gamma(\alpha).\Gamma(\beta)}{\Gamma(\alpha+\beta)},\ \alpha,\beta>0$ وبقسمة طرفي دالة بيتا على $B(\alpha,\beta)$ نحصل على :

$$1=\frac{1}{B(\alpha,\beta)}\int_0^1 x^{\alpha-1}.(1-x)^{\beta-1}\,dx$$

ومما تقدم نستنتج ان X متغير عشوائي بدالة كثافة احتمالية :

$$f(x;\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)}x^{\alpha-1}.(1-x)^{\beta-1}\quad ,0<x<1$$

وفي هذه الحالة يقال ان المتغير العشوائي X يتوزع وفق دالة توزيع بيتا بالمعلمتين α,β . وبالرموز فان $X\sim B(\alpha,\beta)$ واذا كانت $\alpha=\beta=1$ فاننا سوف نحصل على دالة التوزيع المنتظم المستمر على الفترة $(0,1)$. والشكل (٦ ـ ١٨) يبين مخططات لدوال مختلفة لتوزيع بيتا :



الشكل (٦ – ١٨) : الاشكال المختلفة لدالة توزيع بيتا

٦ ـ ٥ ـ ١ : الدالة التوزيعية

تعرف الدالة التوزيعية في توزيع بيتا بانها دالة (او تكامل) بيتا الناقصة Incomplete beta function وهي :

$$F(x) = P_r(X \leq x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x y^{\alpha-1}.(1-y)^{\beta-1}\, dy$$

وهنالك جداول خاصة بهذه الدالة تبين قيم x التي تعطي احتمالا متراكما مقداره $F(x)$ لاحظ احد هذه الجداول (جدول ٥ ملحق ب)الذي صمم لايجاد قيم x التي تجعل $F(x) = 0.05$ عند قيم مختلفة الى $\alpha^* = \alpha - 1, \beta^* = \beta - 1$ فمثلاً قيمة x التي تجعل $F(x) = 0.05$ عندما $\alpha = 4, \beta = 6$ (اي عندما $\alpha^* = 3, \beta^* = 5$) هي $x = 0.16875$ ،

٦ ـ ٥ ـ ٢ : الدالة المولدة للعزوم حول نقطة الاصل

يمتلك توزيع بيتا دالة مولدة للعزوم حول نقطة الاصل ، الا انه من الصعوبة صياغتها بشكل مألوف كما في التوزيعات السابقة التي درسناها، وامكن اشتقاق صيغة لهذه الدالة (على الرغم من كونها صيغة معقدة بعض الشيء) باستخدام ماتسمى بـ « الدالة الزائدية المندمجة Confluent hypergeometric function » وكما هو مبين بالآتي :

$$M_X(t) = Ee^{tX} = M(\alpha; \alpha+\beta; t)$$

$$= \sum_{j=0}^{\infty} \frac{\alpha^{[j]}}{(\alpha+\beta)^{[j]}} \cdot \frac{t^j}{j!}$$

حيث انه وبشكل عام $Z^{[j]} = \prod_{j=0}^{\infty} (Z+j),\ Z^{[0]} = 1$

وهذا يعني أن :

$$M_X(t) = 1 + \frac{\alpha}{\alpha+\beta} \cdot \frac{t}{1!} + \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \cdot \frac{t^2}{2!} + \ldots$$

وبشكل خاص فان العزم الاول والثاني حول نقطة الاصل هما :

وان

$$EX = \mu_x = M'_X(0) = \frac{\alpha}{\alpha + \beta}$$

$$EX^2 = M''_X(0) = \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)}$$

عليه فان التباين في هذا التوزيع هو :

$$\sigma_x^2 = EX^2 - \mu_x^2 = \frac{\alpha\beta}{(\alpha + \beta)^2 . (\alpha + \beta + 1)}$$

وكبديل لهذه الدالة يمكن اشتقاق صيغة للعزم ذي المرتبة r حول نقطة الاصل وكما يلي :

$$EX^r = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^r . x^{\alpha - 1} . (1 - x)^{\beta - 1} dx$$

$$= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(r + \alpha) - 1} . (1 - x)^{\beta - 1} dx$$

$$= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} . \frac{\Gamma(r + \alpha) . \Gamma(\beta)}{\Gamma(r + \alpha + \beta)}$$

فاذن

$$EX^r = M^{(r)}(\alpha ; \alpha + \beta ; 0) = \frac{\Gamma(r + \alpha) . \Gamma(\alpha + \beta)}{\Gamma(\alpha) . \Gamma(r + \alpha + \beta)} , r = 1, 2, \ldots$$

٦ ـ ٥ ـ ٣ : المنوال ونقاط الانقلاب في دالة توزيع بيتا

كماهو معلوم فان المنوال يمثل قيمة x المعرفة في Ω_x الناتجة من حل المعادلة التفاضلية $f'(x) = 0$ بشرط ان $f''(x) < 0$ فاذن

$$f(x) = C . x^{\alpha - 1} . (1 - x)^{\beta - 1} , C = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

او ان

$$\ln f(x) = \ln c + (\alpha - 1)\ln x + (\beta - 1)\ln(1 - x)$$

وباشتقاق الطرفين بالنسبة الى x نحصل على

$$f'(x) = f(x).\left(\frac{\alpha - 1}{x} - \frac{\beta - 1}{1 - x}\right)$$

وبوضع $f'(x) = 0$ فان

$$f(x)\left(\frac{\alpha - 1}{x} - \frac{\beta - 1}{1 - x}\right) = 0$$

وحيث ان $f(x) > 0$ فان $\frac{\alpha - 1}{x} - \frac{\beta - 1}{1 - x} = 0$ وهذا يعني ان

$x = \frac{\alpha - 1}{\alpha + \beta - 2}, \alpha, \beta > 1$، كذلك فان المشتقة الثانية نسبة الى x هي :

$$f''(x) = f(x)\left[-\frac{\alpha - 1}{x^2} - \frac{\beta - 1}{(1 - x)^2} + \left(\frac{\alpha - 1}{x} - \frac{\beta - 1}{1 - x}\right)^2\right]$$

وان

$$f''(x)\Big]_{x = \frac{\alpha - 1}{\alpha + \beta - 2}} = -f(x)\left[\frac{(\alpha + \beta - 2)^2}{\alpha - 1} + \frac{(\alpha + \beta - 2)^2}{\beta - 1}\right]$$

ويلاحظ مايلي :

ان $f''(x) < 0$ عندما $\beta > 1, \alpha > 1$ وذلك يعني ان الدالة $f(x)$ تمتلك نهاية عظمى وان $x = \frac{\alpha - 1}{\alpha + \beta - 2}$ تمثل قيمة المنوال لهذا التوزيع وان $f''(x) > 0$ عندما $\beta < 1, \alpha < 1$ وذلك يعني ان الدالة $f(x)$ تمتلك نهاية صغرى وفي هذه الحالة لايمكن تعريف المنوال في هذا التوزيع . كذلك يمكن بيان ان لمنحنى دالة توزيع بيتا نقطتي انقلاب عندما $\alpha, \beta > 1$ هما

$$\frac{\alpha - 1}{\alpha + \beta - 2} \pm \frac{1}{\alpha + \beta - 2}\left(\frac{(\alpha - 1)(\beta - 1)}{\alpha + \beta - 3}\right)^{\frac{1}{2}}$$

ونترك للقاريء برهنة ذلك . ويلاحظ ان هاتين النقطتين تقعان على بعد متساو الى يمين ويسار المنوال وهما قيمتان حقيقيتان معرفتان في الفترة $(0, 1)$ واذا كانت $\beta = 1, \alpha = 2$ فان $f(x) = 2x$ وهذا يعني ان مخطط الدالة

$f(x)$ عبارة عن خط مستقيم يمر من نقطة الاصل ميله 2 . واذا كانت $\beta = 2, \alpha = 1$ فان $f(x) = 2(1 - x)$ وفي هذه الحالة فان مخطط الدالة $f(x)$ عبارة عن خط مستقيم يقطع المحور الصادي عند النقطة (0 , 2) ميله 2 ـ (لاحظ الشكل ٦ ـ ١٨) .

٦ ـ ٥ ـ ٤ : الالتواء في توزيع بيتا .

حيث ان صيغ الوسط والتباين والمنوال في هذا التوزيع معرفة عليه واعتماداً على هذه الصيغ فان معامل الالتواء في توزيع بيتا هو :

$$S_k = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

لكن

$$\text{Mean} - \text{Mode} = \frac{\alpha}{\alpha + \beta} - \frac{\alpha - 1}{\alpha + \beta - 2}$$

$$= \frac{\beta - \alpha}{(\alpha + \beta)(\alpha + \beta - 2)}$$

فاذن

$$S_k = \frac{(\beta - \alpha)\sqrt{\alpha + \beta + 1}}{(\alpha + \beta - 2)\sqrt{\alpha\beta}} \quad \alpha, \beta > 1$$

ويلاحظ من صيغة معامل الالتواء مايلي :

١ ـ اذا كانت $\alpha = \beta$ فان $S_k = 0$ وذلك يعني ان مخطط الدالة $f(x)$ متماثل حول المحور $x = \text{mode} = \frac{1}{2}$ لاحظ الشكل (٦ ـ ١٨) .

٢ ـ اذا كانت $\beta > \alpha$ فان $S_k > 0$ وذلك يعني ان مخطط الدالة $f(x)$ ملتو التواء موجب ، وتزداد شدة الالتواء بزيادة الفرق بين α, β

٣ ـ اذا كانت $\alpha > \beta$ فان $S_k < 0$ وذلك يعني ان مخطط الدالة $f(x)$ ملتو التواء سالب ، وتزداد شدة الالتواء بزيادة الفرق بين α, β

٦ ـ ٥ ـ ٥ : حالات خاصة لتوزيع بيتا

نستعرض في هذه الفقرة بعض الحالات الخاصة لتوزيع بيتا التي يتم الاستفادة منها في نظرية (المسارات العشوائية "Randomwalks وهي :

١ ـ توزيع الجيب القوسي Arc - sine distribution

في حالة اختيار $\alpha = \beta = \frac{1}{2}$ عندئذ فان

$$f(x) = \frac{1}{\pi} x^{-\frac{1}{2}} . (1-x)^{-\frac{1}{2}} \; ; 0 < x < 1$$

ان تسمية هذا التوزيع بـ « توزيع الجيب القوسي » ناجمة عن كون ان :

$$F(x) = P_r(X \le x) = \frac{2}{\pi} \sin^{-1} \sqrt{x}$$

ومخطط هذا التوزيع موضح في الشكل (٦ ـ ١٨) .

٢ ـ توزيع الجيب القوسي العمومي Generalized arc-sine dist.

اذا كانت $\alpha + \beta = 1$ وان $\alpha \neq \beta$ عندئذٍ نحصل على حالة خاصة من توزيع بيتا هي توزيع الجيب القوسي العمومي . فمثلاً اذا كانت $\alpha = \frac{1}{4}$ ، $\beta = \frac{3}{4}$ فان

$$f(x) = \frac{1}{\Gamma\left(\frac{1}{4}\right) . \Gamma\left(\frac{3}{4}\right)} x^{-\frac{3}{4}} . (1-x)^{-\frac{1}{4}} \; ; 0 < x < 1$$

٦ ـ ٥ ـ ٦ : امثلة

مثال (١) : اذا كان$X \sim B(2,3)$فان :

$$1 - f(x) = 12x(1-x)^2 \; ; 0 < x < 1$$

$$2 - F(x) = 12\int_0^x u(1-u)^2 \, du$$

$$= 6x^2 - 8x^3 + 3x^4 \; ; 0 < x < 1$$

لاحظ ان

$$F(0) = 0, F(1) = 1, F(0{\cdot}05) = 0{\cdot}6875$$

$$3 - \mu_x = \frac{\alpha}{\alpha+\beta} = \frac{2}{5}$$

$$\sigma_x^2 = \frac{\alpha\beta}{(\alpha+\beta)^2 . (\alpha+\beta+1)} = \frac{1}{25}$$

$$4 - M_X(t) = M(2;5;t)$$

مثال (٢) : ليكن $X \sim B(\alpha,\beta)$. جد الدالة المولدةلعزوم $Y = \ln \dfrac{X}{1-X}$.

الحل :

افرض ان المتغير Y يمتلك دالة مولدة للعزوم حول نقطة الاصل . وهذا يعني ان

$$M_Y(t) = Ee^{tY} = Ee^{t \ln \frac{X}{1-X}} \quad = Ee^{\ln\left(\frac{X}{1-X}\right)^t}$$

$$\therefore M_Y(t) = E\left(\frac{X}{1-X}\right)^t$$

وحيث ان $X \sim B(\alpha, \beta)$ فاذن :

$$E\left(\frac{X}{1-X}\right)^t = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 \frac{x^t}{(1-x)^t} . x^{\alpha-1}\ (1-x)^{\beta-1}\, dx$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{(\alpha+t)-1} . (1-x)^{(\beta-t)-1}\, dx$$

ويلاحظ ان التكامل اعلاه يمثل تكامل بيتا بالمعلمتين $(\alpha+t), (\beta-t)$ وان قيمة هذا التكامل هي :

$$\frac{\Gamma(\alpha+t) . \Gamma(\beta-t)}{\Gamma(\alpha+\beta)}$$

عليه فان

$$M_Y(t) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} . \frac{\Gamma(\alpha+t).\Gamma(\beta-t)}{\Gamma(\alpha+\beta)} = \frac{\Gamma(\alpha+t).\Gamma(\beta-t)}{\Gamma(\alpha) . \Gamma(\beta)}, \beta > t$$

تمارين عن توزيع بيتا

٦ ـ ٢٢ : اذا كان $X \sim B(\alpha, \beta)$. اشتق صيغة لكل ممايلي :

أ ـ الوسط التوافقي

ب ـ معامل الاختلاف

٦ ـ ٢٣ : لكل حالة من الحالات التالية جد قيمة الثابت c بحيث ان $f(x)$ هي دالة لتوزيع بيتا

أ ـ $f(x) = c . x^2 (1-x)^5$

ب ـ $f(x) = c . x^3 (1-x)^3$

جـ ـ $f(x) = c(x - x^2)^{0.5}$

٦ ـ ٢٤ : جد قيمة الثابت c بحيث ان $f(x) = cx(5-x)^6, 0 < x < 5$ هي دالة لتوزيع بيتا .

٦ ـ ٢٥ : بفرض ان $X \sim B(\alpha, \beta)$. جد $E(1-x)^\alpha, EX^\beta, EX^\beta . (1-X)^\alpha$

٦ ـ ٢٦ : اذا كان $X \sim B(\alpha, \beta)$. برهن ان $Y = 1 - X \sim B(\beta, \alpha)$

٦ ـ ٦ : توزيعات مستمرة اخرى
Other continuous distributions

سوف نستعرض في هذه الفقرة بعض التوزيعات المستمرة الاخرى . غير المذكورة في الفقرات السابقة . ذات العلاقة بالنظرية الاحصائية من خلال عرض لاهم خصائص هذه التوزيعات دون اللجوء الى البراهين والاشتقاقات اللازمة لهذه الخصائص (ماعدا في بعض الحالات التي تتطلب التوضيح فقط) بهدف اطلاع القاريء بها .

٦ ـ ٦ ـ ١ : توزيع كوشي Cauchy's distribution

ينسب هذا التوزيع للعالم الرياضي الفرنسي A . L . Cauchy الذي تمكن من اشتقاق هذا التوزيع ونشره عام ١٨٥٣ . ويعرف هذا التوزيع على النحو التالي : يقال ان المتغير العشوائي X يتوزع وفق توزيع كوشي اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

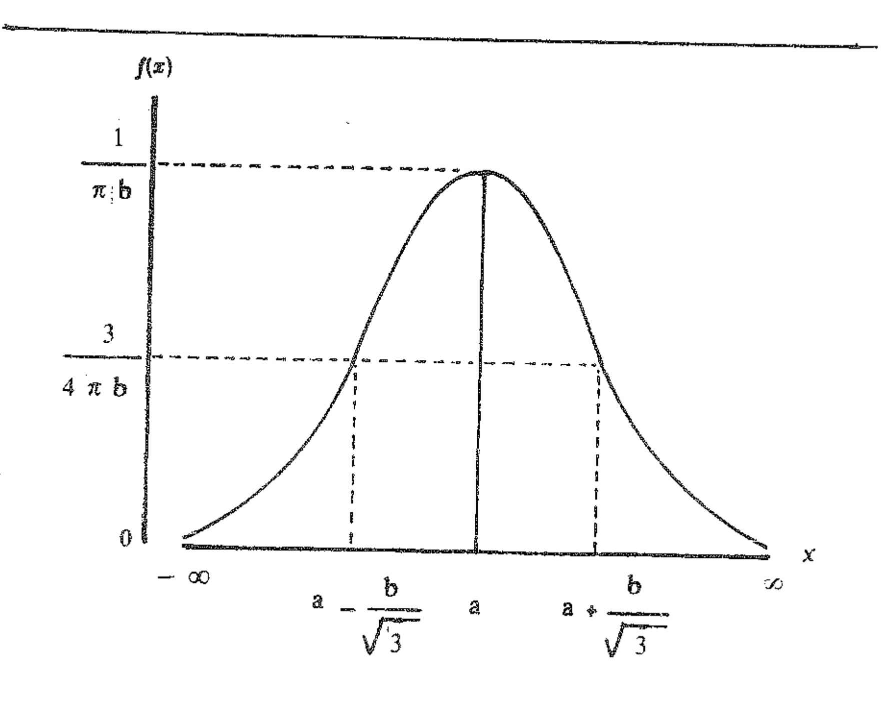
$$f(x\,;a\,,b) = \frac{b}{\pi\left[\,b^2 + (x-a)^2\,\right]}\;;\; -\infty < x < \infty$$

حيث a , b معلمتا التوزيع . $b > 0\,,\ -\infty < a < \infty$. وان a تمثل مقياس لموقع التوزيع على المحور السيني . b مقياس لتشتت التوزيع . واذا كانت $b = 1\,,\ a = 0$ عندئذٍ نحصل على توزيع كوشي المعياري المعطاة دالة كثافته الاحتمالية بـ : $f(x) = \dfrac{1}{\pi(1+x^2)}\;;\; -\infty < x < \infty$. وفيما يلي بعض خصائص هذا التوزيع :

١ ـ ان منحنى دالة توزيع كوشي متماثل حول المحور $x = a$. وهذا يعني ان المنوال في توزيع كوشي هو قيمة المعلمة a وان لاي عدد موجب مثل h فان $f(a-h) = f(a+h)$. وبذلك فان الدالة $f(x)$ تكون في نهايتها العظمى عندما $x = a$ وقيمة الدالة عند هذه النقطة هي :

$$\text{Max.}\, f(x) = f(x)\big]_{x=a} = \frac{1}{\pi b}$$

ويلاحظ من هذه الصيغة ان منحنى دالة هذا التوزيع يزداد تفلطحاً عند زيادة قيمة b . كذلك فان لمنحنى دالة هذا التوزيع نقطتي انقلاب هما $x = a \pm \frac{b}{\sqrt{3}}$ وان قيمة الدالة عند هاتين النقطتين هي $\frac{3}{4\pi b}$. والشكل (٦ ـ ١٩) يوضح مخطط دالة هذا التوزيع :



الشكل (٦ ـ ١٩) : مخطط دالة توزيع كوشي

٢ ـ الدالة التوزيعية لتوزيع كوشي هي :

$$F(x) = P_r(X \leq x) = \frac{b}{\pi} \int_{-\infty}^{x} \frac{du}{[b^2 + (u - a)^2]}$$

$$= \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left(\frac{x - a}{b} \right)$$

ونظراً لامكانية حساب الاحتمال المتراكم بشكل مباشر من خلال $F(x)$ فان ذلك لا يستدعي تكوين جداول خاصة بهذا التوزيع . فمثلاً اذا كانت $a = 4$, $b = 6$ فان $F(6) = 0.6024163$ وان $F(1) = 0.3524164$. كذلك يمكن ملاحظة ان $F(a) = \frac{1}{2}$ وهذا يعني ان الوسيط في هذا التوزيع هو قيمة المعلمة a ، وان الربيع الاول فيه هو قيمة x الناتجة من حل $F(x) = \frac{1}{4}$ وهي $Q_1 = a - b$. حسب خاصية التماثل فان قيمة الربيع الثالث هي $Q_3 = a + b$ او انها قيمة x الناتجة من حل $F(x) = \frac{3}{4}$.

٣ ـ ان العزم ذو المرتبة r حول نقطة الاصل غير موجود . وذلك يعني عدم امكانية تحديد متوسط التوزيع وتباينه وكذلك اي عزم من عزومه . وبهدف توضيح ذلك افرض ان $b = 1, a = 0$ فان :

$$EX^r = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x^r}{1 + x^2} dx$$

وباستخدام التكامل بالتجزئة من خلال الفرض بان

$$u = x^{r-1} \rightarrow du = (r-1)x^{r-2}dx, dv = \frac{x}{1 + x^2}dx \rightarrow v = \frac{1}{2}\ln(1 + x^2)$$

فان :

$$EX^r = \frac{1}{2\pi} x^{r-1}.\ln(1 + x^2)\Big]_{-\infty} - \frac{r-1}{2\pi} \int_{-\infty}^{\infty} x^{r-2}.\ln(1 + x^2)dx$$

لاحظ ان ناتج التكامل متباعد وهذا يعني ان EX^r غير موجود . ولان هذا التوزيع متماثل عند النقطة $x = a$ فانه يمكن القول ان متوسط هذا التوزيع هو قيمة المعلمة a . كما ويمكن اعتبار مربع الانحراف الربيعي كمقياس بديل للتباين في هذا التوزيع حيث ان الانحراف الربيعي في هذا التوزيع هو :

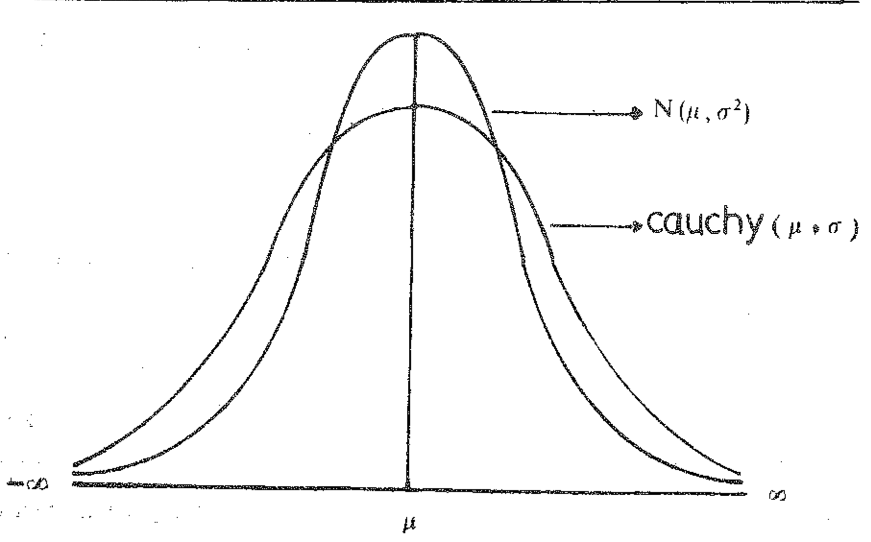
$$Q = \frac{Q_3 - Q_1}{2} = \frac{(a + b) - (a - b)}{2} = b$$

٤ ـ الدالة المولدة لعزوم توزيع كوشي غير موجودة بسبب ان $\int_{-\infty}^{\infty} e^{tx} f(x)\, dx$ متباعد .

٥ ـ الدالة المميزة لتوزيع كوشي هي :

$$\phi_x(t) = Ee^{itX} = e^{ita - b|t|}$$

٦ ـ بشكل عام فان منحنى دالة توزيع كوشي اكثر تفلطحاً من منحنى دالة $N(\mu, \sigma^2)$ عند الفرض بان $b = \sigma , a = \mu$ اي ان المتغير العشوائي X يتوزع وفق دالة توزيع كوشي بالمعلمتين σ, μ . وكما هو موضح في الشكل (٦ ـ ٢٠)



الشكل (٦ ـ ٢٠) ، مقارنة بين توزيع طبيعي وتوزيع كوشي

٧ ـ اذا كانت $X_1, X_2, \ldots, X_n$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim \text{cauchy}(a_i, b_i)$ وان $Y = \sum_{i=1}^{n} C_i X_i$. حيث C_i ثوابت حقيقية . عندئذ

$$Y \sim \text{cauchy}\left(\sum_{i=1}^{n} c_i a_i , \sum_{i=1}^{n} |c_i| b_i \right).$$

ويمكن استنتاج هذه الخاصية (خاصية الجمع) من خلال الدالة المميزة .

وبشكل خاص اذا كانت $X_1, X_2, \dots, X_n$ تمثل قياسات عينة عشوائية مسحوبة من توزيع كوشي بالمعلمتين b, a فان الوسط الحسابي $\bar{X}$ لهذه العينة سوف يتوزع كتوزيع كوشي بنفس المعلمتين b , a اي ان $\bar{X} \sim \text{cauchy}(a, b)$ ويمكن اثبات ذلك من خلال وضع

$$a_i = a , b_i = b , C_i = \frac{1}{n} \quad \forall_i = 1, 2, \dots, n$$

٨ ـ اذا كان $Z_1 \sim N(0, 1)$ مستقل عن $Z_2 \sim N(0, 1)$ عندئذٍ فان $X = \frac{Z_1}{Z_2} \sim \text{cauchy}(a = 0, b = 1)$, اي ان النسبة ما بين متغيرين عشوائيين طبيعيين معياريين تتوزع كتوزيع كوشي معياري .

٦ ـ ٦ ـ ٢ : التوزيع اللوغارتمي الطبيعي

The Log - Normal distribution

ان للتوزيع اللوغارتمي الطبيعي اهمية لاتقل عن اهمية التوزيع الطبيعي في الجوانب التطبيقية للنظرية الاحصائية . فهو احد التوزيعات الهامة التي تدخل في موضوع الرقابة على جودة الانتاج ، وكذلك في الدراسات المتعلقة بعلم الحشرات entomological problems والكيمياء الجيولوجية geo – chemistry وغيرها من الموضوعات التي يدخل فيها الاحصاء كأداء للتحليل . ويعرف هذا التوزيع على النحو التالي ، اذا كان $Y \sim N(\mu, \sigma^2)$ وان $X = e^Y$ عندئذٍ يقال ان X يتوزع كتوزيع لوغارتمي طبيعي بالمعلمتين μ, σ^2 .

البرهان :

ان

$$X = e^Y \rightarrow Y = \ln X , dy = \frac{dx}{x} , 0 < x < \infty$$

الان

$$\frac{1}{\sqrt{2\pi}\ \sigma} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy = 1, Y \sim N(\mu, \sigma^2)$$

فاذن

$$\frac{1}{\sqrt{2\pi}\ \sigma} \int_{0}^{\infty} e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2} . \frac{dx}{x} = 1$$

وهذا يعني ان دالة الكثافة الاحتمالية للمتغير X هي :

$$f(x; \mu, \sigma^2) = \frac{1}{x\sqrt{2\pi}\ \sigma} . e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2} ; x > 0$$

حيث ان $-\infty < \mu < \infty$, $\sigma > 0$. وبالرموز فان $X \sim \log N(\mu, \sigma^2)$.

وفيما يلي بعض خصائص هذا التوزيع .

١ ــ اذا كان $X \sim \log N(\mu, \sigma^2)$ فان $Z = \frac{\ln x - \mu}{\sigma} \sim N(0, 1)$

٢ ــ ان الدالة التوزيعية في التوزيع اللوغارتمي الطبيعي هي :

$$\begin{aligned} F(x) &= P_r(X \le x) = P_r(\ln X \le \ln x) \\ &= P_r(Y \le \ln x) \quad ; Y \sim N(\mu, \sigma^2) \\ &= P_r\left(\frac{Y-\mu}{\sigma} \le \frac{\ln x - \mu}{\sigma}\right) \\ &= P_r\left(Z \le \frac{\ln x - \mu}{\sigma}\right) \quad ; Z \sim N(0, 1) \\ &= \Phi\left(\frac{\ln x - \mu}{\sigma}\right) \end{aligned}$$

حيث ان $\Phi(\cdot)$ تعني الدالة التوزيعية للتوزيع الطبيعي المعياري . وهذا يعني انه يمكن استخدام جداول التوزيع الطبيعي المعياري في حساب التراكم

الاحتمالي للتوزيع الطبيعي اللوغارتمي . فمثلاً اذا كان $X \sim \log N(2, 4)$ فان :

$$F(0) = \Phi(-\infty) = 0, F(1) = \Phi(-1) = 0{\cdot}1587, F(2) = \Phi(-0{\cdot}65) = 0{\cdot}2578$$

٣ ـ ان العزم ذو المرتبة r حول نقطة الاصل لتوزيع $\log N(\mu, \sigma^2)$ هو :

$$EX^r = E(e^Y)^r = Ee^{rY} = M_Y(r) \quad ; Y \sim N(\mu, \sigma^2)$$

$$= e^{\mu r + \frac{1}{2} r^2 \sigma^2} \quad , r = 1, 2, \ldots$$

واضح ان

$$\mu_x = EX = e^{\mu + \frac{1}{2}\sigma^2}$$

وان

$$EX^2 = e^{2\mu + 2\sigma^2}$$

وبذلك فان

$$V(X) = \sigma_x^2 = EX^2 - (EX)^2$$

$$= e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2} = e^{2\mu}(e^{2\sigma^2} - e^{\sigma^2})$$

٤ ـ اذا كانت $X_1, X_2, \ldots, X_n$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim \log N(\mu, \sigma^2)$ عندئذٍ $V = \prod_{i=1}^{n} X_i \sim \log N(\mu, n\sigma^2)$.

٥ ـ اذا كان $X \sim \log N(\mu, \sigma^2)$ عندئذٍ $E(\ln X) = \mu$, $V(\ln X) = \sigma^2$.

٦ ـ ان المنوال في التوزيع اللوغارتمي الطبيعي هو قيمة x الناتجة من حل المعادلة التفاضلية $f'(x) = 0$ بشرط ان $f''(x) < 0$ وهي $x = e^{\mu - \sigma^2}$.

وعليه فان

$$\text{Max.}\, f(x) = \left(\sqrt{2\pi} \,.\, \sigma .\, e^{\mu - \frac{1}{2}\sigma^2} \right)^{-1}$$

كذلك يمكن بيان ان لمنحنى دالة هذا التوزيع نقطتي انقلاب ناتجتين من حل $f''(x) = 0$ بشرط ان $f'''(x) \neq 0$. هاتين النقطتين هما :

$$x = e^{\alpha} \quad , \quad \alpha = \left(\mu - \frac{3\sigma^2}{2}\right) \pm \sigma \sqrt{1 + \frac{1}{4}\sigma^2}$$

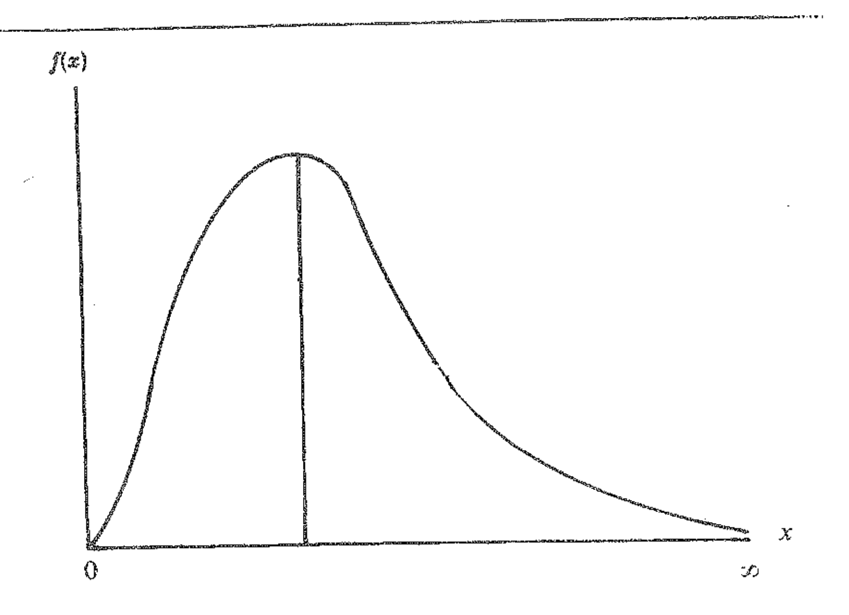
٧ ـ ان الوسيط لهذا التوزيع يمثل قيمة x التي تحقق $F(x) = \frac{1}{2}$ او $\Phi\left(\frac{\ln x - \mu}{\sigma}\right) = \frac{1}{2}$. لكن كما هو معلوم فان $\Phi(0) = \frac{1}{2}$ فاذن

$$\frac{\ln x - \mu}{\sigma} = 0 \rightarrow \ln x = \mu \quad \therefore x = e^{\mu}$$

ويتضح مما تقدم ان متوسط التوزيع اكبر من وسيطه وهذا اكبر من منوال التوزيع اي ان

$$\mu_x = e^{\mu + \frac{1}{2}\sigma^2} > \text{median} = e^{\mu} > \text{mode} = e^{\mu - \sigma^2}$$

وهذا يعني ان منحنى دالة هذا التوزيع هو ذو التواء موجب . والشكل (٦ ـ ٢١) . يوضح مخطط دالة هذا التوزيع :



الشكل (٦ ـ ٢١) : مخطط لدالة توزيع لوغارتمي طبيعي .

٦ ـ ٦ ـ ٣ : التوزيع السوقي (اللوجستي) **Logistic distribution**

تبرز استخدامات هذا التوزيع وبشكل خاص في الدراسات المتعلقة بعلوم الحياة Biological assay والعلوم الزراعية والطبية وبشكل عام في الدراسات ذات الطابع التجريبي . وفيما يلي تعريف لهذا التوزيع : يقال ان المتغير العشوائي X هو ذو توزيع سوقي اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

$$f(x;\alpha,\beta) = \frac{1}{4\beta}.\operatorname{sech}^2\left[\frac{1}{2}\left(\frac{x-\alpha}{\beta}\right)\right]$$
$$= \frac{1}{\beta}[e^{-\left(\frac{x-\alpha}{\beta}\right)}].[1+e^{-\left(\frac{x-\alpha}{\beta}\right)}]^{-2} ; -\infty < x < \infty$$

حيث α , β تمثلان معلمتي التوزيع وان $\beta > 0 , -\infty < \alpha < \infty$ وبالرموز فان $X \sim \text{Logistic}(\alpha,\beta)$. وفيما يلي بعض خصائص هذا التوزيع :

١ ـ ان الدالة التوزيعية لهذا التوزيع هي :

$$F(x) = P_r(X \leq x) = \frac{1}{2}\left[1 + \tanh\left[\frac{1}{2}\left(\frac{x-\alpha}{\beta}\right)\right]\right]$$
$$= [1 + e^{-\left(\frac{x-\alpha}{\beta}\right)}]^{-1}$$

كذلك يمكن التعبير عن $f(x)$ بدلالة $F(x)$ من خلال العلاقة التالية $f(x) = \frac{1}{\beta}F(x)(1-F(x))$ ونترك برهنة ذلك للقارىء .

٢ ـ ان الدالة المولده لعزوم X حول نقطة الاصل هي :

$$M_X(t) = e^{\alpha t}.\Gamma(1-\beta t).\Gamma(1+\beta t)$$
$$= e^{\alpha t}.(\pi\beta t).\csc(\pi\beta t)$$

ومن خلال هذه الدالة يمكن بيان ان .

$$\mu_x = EX = M'_x(0) = \alpha, EX^2 = M''_x(0) = \alpha^2 + \frac{\pi^2\beta^2}{3}$$

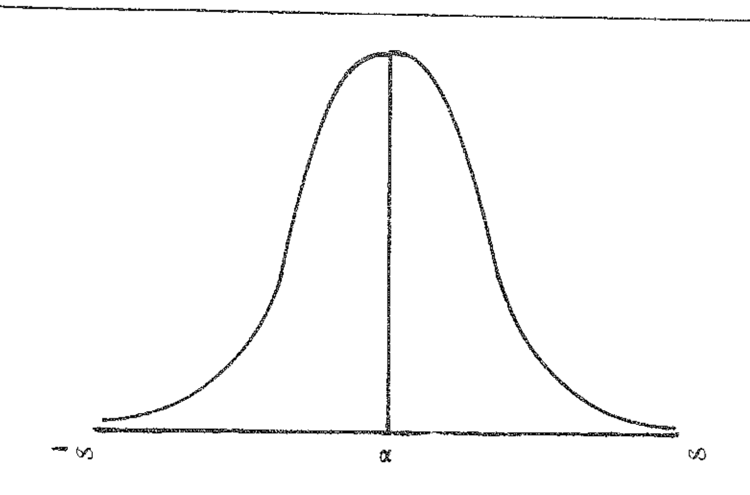
وان

$$\sigma_x^2 = \frac{\pi^2\beta^2}{3}$$

٣ ـ يتحقق المنوال في هذا التوزيع عندما $x = \alpha$ وهذا ناتج من حل المعادلة التفاضلية $f'(x) = 0$ بشرط ان $f''(x) < 0$. وعندئذٍ فان

$$\text{Max}.\, f(x) = f(x)\Big]_{x=\alpha} = \frac{1}{4\beta}$$

٤ ـ كذلك يتحقق الوسيط في هذا التوزيع عندما $x = \alpha$ وهذا ناتج من حل الصيغة $F(x) = \frac{1}{2}$. ونلاحظ في هذا التوزيع تساوي الاوساط الثلاثة (الوسط . الوسيط . المنوال) وهذا يعني ان منحنى دالة هذا التوزيع متماثل حول المحور $x = \alpha$. والشكل (٦ ـ ٢٢) يوضح مخطط هذا التوزيع .



الشكل (٦ ـ ٢٢) . مخطط لدالة التوزيع السوقي .

٤ ـ اذا كانت $\beta = 1 , \alpha = 0$ عندئذٍ نحصل على الشكل المعياري لهذا التوزيع . وفي هذه الحالة تكون :

$$f(x) = e^{-x}(1 + e^{-x})^{-2} = \frac{1}{4} \operatorname{sech}^2 \frac{x}{2} \; ; - \infty < x < \infty$$

وان

$$F(x) = [1 + e^{-x}]^{-1} = \frac{1}{2}\left[1 + \tanh \frac{x}{2}\right]$$

كذلك فان

$$M_x(t) = \Gamma(1 - t) . \Gamma(1 + t) = \pi t . \csc \pi t$$

وان

$$\mu_x = 0 , \sigma_x^2 = \frac{\pi^2}{3}$$

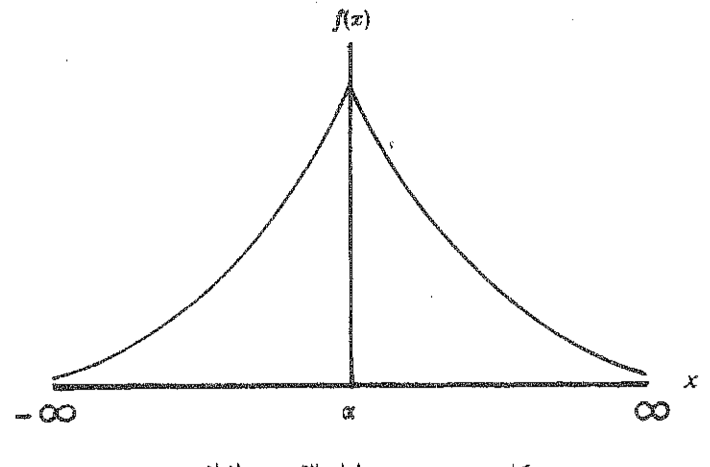
٦ ـ ٦ ـ ٤ : توزيع لاپلاس Laplace distribution

يعتبر العالم الفرنسي Laplace اول من اكتشف هذا التوزيع وكان ذلك عام ١٧٧٤ . ويعرف هذا التوزيع على النحو الآتي : يقال ان المتغير العشوائي X يتوزع كتوزيع لاپلاس اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

$$f(x, \alpha, \beta) = \frac{1}{2\beta} e^{- \frac{|x - \alpha|}{\beta}} \quad ; - \infty < x < \infty$$

حيث ان α , β تمثلان معلمتي التوزيع بحيث ان $\beta > 0 , - \infty < \alpha < \infty$ وبالرموز فان $X \sim Lap(\alpha , \beta)$. والشكل (٦ ـ ٢٣) يوضح مخطط دالة هذا التوزيع :

ويلاحظ من الشكل (٦ ـ ٢٣) ان الدالة $f(x)$ تكون في نهايتها العظمى عندما $x = \alpha$ (اي ان المنوال في هذا التوزيع هو قيمة المعلمة α) وان $Max. f(x) = f(x)]_{x=\alpha} = \frac{1}{2\beta}$ كذلك فان منحنى دالة هذا التوزيع متماثل حول المحور $x = \alpha$ وهذا يعني انه لأي عدد موجب مثل h فان $f(\alpha - h) = f(\alpha + h)$. وفيما يلي بعض خصائص هذا التوزيع :



شكل (٦ ــ ٢٣) ، مخطط دالة توزيع لاپلاس

١ ــ الدالة التوزيعية في توزيع لاپلاس هي :

$$F(x) = \frac{1}{2} e^{-\left(\frac{\alpha - x}{\beta}\right)} \quad , x \leq \alpha$$

$$= 1 - \frac{1}{2} e^{-\left(\frac{x - \alpha}{\beta}\right)} \quad , x \geq \alpha$$

وحيث انه من السهولة حساب قيم $F(x)$ مباشرة من هذه الدالة لذا فانه ليس من الضروري اعداد جداول خاصة بهذا التوزيع .

٢ ــ الدالة المولدة لعزوم توزيع لاپلاس حول نقطة الاصل هي :

$$M_X(t) = \frac{e^{\alpha t}}{1 - \beta^2 t^2} \; ; t < \frac{1}{\beta}$$

وان الدالة المولدة التراكمية هي

$$K_X(t) = \ln M_X(t) = \alpha t - \ln(1 - \beta^2 t^2)$$

ويتضح من هذه الدالة ان :

$$\mu_x = K'_X(0) = \alpha , \sigma_x^2 = K''_X(0) = 2\beta^2$$

٣ ـ اذا كانت $\beta = 1 , \alpha = 0$ عندئذٍ نحصل على الشكل المعياري لتوزيع لابلاس . وفي هذه الحالة فان :

$$f(x) = \frac{1}{2} e^{-|x|} \quad , -\infty < x < \infty$$

وان

$$F(x) = \frac{1}{2} e^{x} \qquad , x \leq 0$$

$$= 1 - \frac{1}{2} e^{-x} , x \geq 0$$

كذلك فان

$$M_X(t) = (1 - t^2)^{-1} , t < 1$$

٦ ـ ٦ ـ ٥ : توزيع وايبل Weibull distribution

ينسب هذا التوزيع الى الفيزيائي السويدي **Waloddi Weibull** الذي اشتق واستخدم هذا التوزيع عام ١٩٣٩ في دراسة خصائص العدد المنتجة صناعياً . كذلك استعرض kao عام ١٩٥٨ استخدامات هذا التوزيع في دراسات المعولية reliability studies . كما ان لهذا التوزيع استخدامات هامة في موضوع الرقابة على الجوده quality control . وفيما يلي تعريف لهذا التوزيع: يقال ان المتغير العشوائي X يتوزع وفق دالة توزيع وايبل اذا كانت دالة الكثافة الاحتمالية لهذا المتغير تأخذ الشكل التالي :

$$f(x, a, b) = abx^{b-1} . e^{-ax^b} , x > 0$$

حيث a , b تمثلان معلمتي التوزيع وان $a,b > 0$. ويتضح من هذه الدالة انه اذا كانت $b = 1$ فان توزيع وايبل يختزل الى التوزيع الاسي بالمعلمة a .

ونعرض فيما يلي بعض خصائص هذا التوزيع :

١ ـ الدالة التوزيعية لتوزيع وايبل هي :

$$F(x) = 1 - e^{-ax^{b}}, x > 0$$

ويلاحظ من هذه الدالة ان الامر لايستوجب تكوين جداول خاصة بهذا التوزيع طالما ان مسألة حساب قيم $F(x)$ سهلة جداً من خلال التعويض المباشر عن قيم x

٢ ـ ان العزم ذو المرتبة r حول نقطة الاصل هو :

$$EX^{r} = a^{-rb^{-1}} . \Gamma(rb^{-1} + 1) \quad , r = 1, 2, \ldots$$

ويتضح من هذه الصيغة ان :

$$\mu_x = EX = a^{-b^{-1}} . \Gamma(b^{-1} + 1)$$

وان

$$EX^2 = a^{-2b^{-1}} . \Gamma(2b^{-1} + 1)$$

وهذا يعني ان

$$\sigma_x^2 = a^{-2b^{-1}} . [\Gamma(2b^{-1} + 1) - \Gamma^2(b^{-1} + 1)]$$

٣ ـ المنوال في توزيع وايبل هو قيمة x الناتجة من حل المعادلة التفاضلية $f'(x) = 0$ بحيث ان $f''(x) < 0$ وهي :

$$x = \left(\frac{b - 1}{ab} \right)^{b^{-1}}$$

وبذلك فان

$$Max . f(x) = ab\left(\frac{b-1}{ab}\right)^{1-b^{-1}} . e^{-a\left(\frac{b-1}{ab}\right)}$$

كما وان لمنحنى دالة هذا التوزيع نقطتي انقلاب هما .

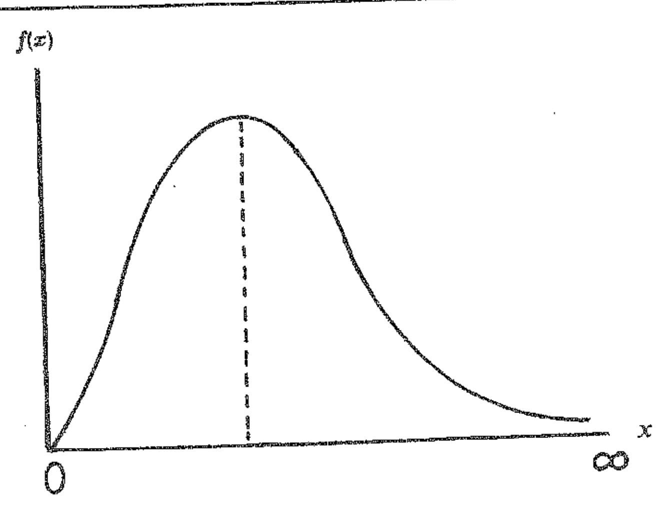
$$x = \left[\frac{3(b-1) \pm \sqrt{(5b-1)(b-1)}}{2ab}\right]^{b^{-1}}$$

٤ ـ الوسيط في توزيع وايبل يمثل قيمة x الناتجة من حل الصيغة $F(x) = \frac{1}{2}$

وهي :

$$x = \left(\frac{\ln(2)}{a}\right)^{b^{-1}}$$

٥ ـ بشكل عام فان منحنى دالة هذا التوزيع ملتو التواء موجب ، وشكل هذاالمنحنى موضح في الشكل (٦ ـ ٢٤) :



الشكل (٦ ـ ٢٤) ، مخطط لدالة توزيع وايبل

٦ ـ ٦ ـ ٦ : توزيع باريتو Pareto distribution

ينسب هذا التوزيع الى العالم الاقتصادي الايطالي Vilfredo Pareto (١٨٤٨ ـ ١٩٢٣) الذي وضع اسس هذا التوزيع . وعلى الرغم من قلة استخدامات هذا التوزيع الا انه لاقى مجالاً كبيراً للتطبيق وخصوصاً في موضوع الاقتصاد من خلال دراسة توزيع الدخول Incomes عندما تكون متجاوزة لحد معلوم مثل a . ويعرف هذا التوزيع على النحو الآتي :

يقال ان التغير العشوائي X يتوزع وفق دالة توزيع باريتو اذا كانت دالة الكثافة الاحتمالية للمتغير X تأخذ الشكل التالي :

$$f(x) = f(x; a, b) = \frac{b}{a}\left(\frac{a}{x}\right)^{b+1}, x \geq a$$

حيث a , b تمثلان معلمتي التوزيع بحيث ان $b > 0, a > 0$. ويتضح من هذه الدالة ان اعظم قيمة لها تتحقق عندما $x = a$ وهي $f(a) = \frac{b}{a}$.

وفيما يلي بعض خصائص هذا التوزيع :

١ ـ الدالة التوزيعية في توزيع باريتوهي :

$$F(x) = 1 - \left(\frac{a}{x}\right)^{b}, x \geq a$$

ومنها يتضح ان $F(\infty) = 1, F(a) = 0$ كذلك لايستدعي الامر اعداد جداول خاصة بهذه الدالة نظراً لسهولة حساب قيمها مباشرة عند معرفتنا بقيمة b , a.

٢ ـ العزم ذو المرتبة r حول نقطة الاصل هو :

$$EX^r = \frac{ba^r}{b - r}, b > r, r = 1, 2, \ldots$$

ويتضح من هذه الصيغة ان :

$$EX = \mu_x = \frac{ab}{b-1}, EX^2 = \frac{a^2 b}{b-2}$$

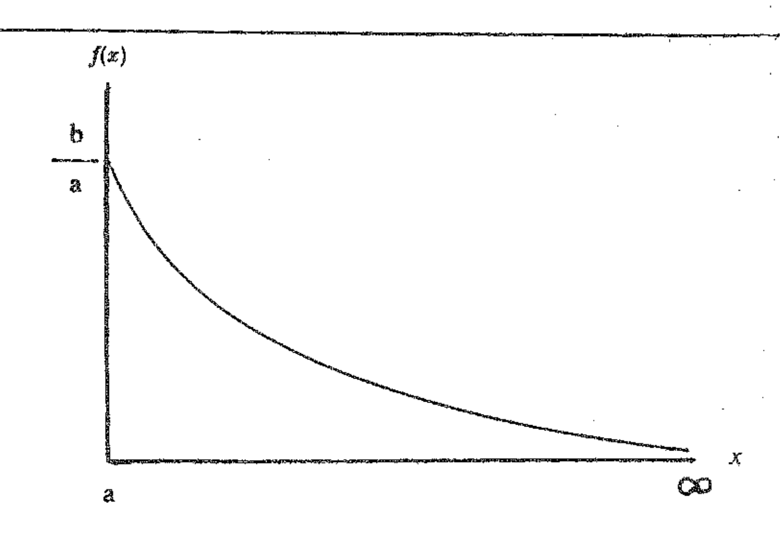
وبذلك فان :

$$\sigma_x^2 = EX^2 - \mu_x^2 = \frac{a^2 b}{(b-2)(b-1)^2}, b > 2$$

٣ ـ الوسيط لتوزيع باريتو هو $x = a2^{b^{-1}}$ بحيث ان

$$F(a2^{b^{-1}}) = \frac{1}{2}$$

والشكل (٦ ـ ٢٥) يوضح مخطط لدالة هذا التوزيع :



الشكل (٦ ـ ٢٥) : مخطط لدالة توزيع باريتو .

٦ ـ ٦ ـ ٧ : توزيع كامبل Gumbel distribution

ويسمى في بعض الاحيان « توزيع القيمة المتطرفة **extreme - value distribution** . يقال ان المتغير العشوائي X يتوزع وفق دالة توزيع كامبل اذا كانت دالة الكثافة الاحتمالية للمتغير X هي :

$$f(x; a, b) = \frac{1}{b} . \exp\left\{ -\left(\frac{x-a}{b}\right) + \exp\left(-\frac{x-a}{b}\right)\right\}$$

حيث exp تعني الرفع للعدد الطبيعي e وان $-\infty < x < \infty$, a, b تمثلان معلمتي التوزيع بحيث ان $-\infty < a < \infty$, $b > 0$. وفيما يلي بعض خصائص هذا التوزيع .

١ ـ ان الدالة التوزيعية في توزيع كامبل هي :

$$F(x) = \exp\left\{-\exp\left(-\frac{x-a}{b}\right)\right\}, -\infty < x < \infty$$

ويتضح من هذه الدالة ان

$$\lim_{x \to -\infty} F(x) = 0 , \lim_{x \to \infty} F(x) = 1$$

كذلك يلاحظ انه لا حاجة لعمل جداول خاصة بهذا التوزيع نظراً لامكانية حساب قيم $F(x)$ مباشرة عند معرفتنا بقيمتي b, a . كذلك يمكن التعبير عن $f(x)$ بدلالة $F(x)$ وفق العلاقة الاتية :

$$f(x) = \frac{1}{b} F(x) . \ln [F(x)]^{-1}$$

ونترك برهنة ذلك للقارىء

٢ ـ ان الدالة المولدة لعزوم توزيع كامبل حول نقطة الاصل هي :

$$M_X(t) = e^{at} . \Gamma(1 - bt) ; t < \frac{1}{b}$$

وان الدالة المولدة التراكمية هي

$$K_X(t) = \ln M_X(t) = at + \ln \Gamma(1 - bt)$$

كذلك فان

$$K'_X(t) = a + \frac{\Gamma'(1 - bt)}{\Gamma(1 - bt)} = a + \psi(1 - bt)$$

حيث ان الدالة ψ تدعى بـ « دالة كاما المضاعفة digamma function » عليه فان

$$EX = K'_X(0) = a + b\psi(1) = a + \lambda b$$

حيث ان $\lambda \simeq 0{\cdot}577216$ يسمى ثابت اويلر Euler's Constant

كذلك فان

$$\sigma_x^2 = K''_x(0) = b^2\psi'(1) = \frac{\pi^2 b^2}{6}$$

وبشكل عام فان

$$K_X^{(r)}(0) = (-b)^r \cdot \psi^{(r-1)}(1) \; ; \; r \geq 2$$

٣ ـ ان المنوال في توزيع كامبل يتحقق عندما $x = a$ وعندئذ فان

$$\text{Max} \, . \, f(x) = f(x)\big]_{x=a} = \frac{1}{be}$$

كما وان لمنحنى دالة هذا التوزيع نقطتي انقلاب هما :

$$x = a \pm 0{\cdot}9624236 \, b$$

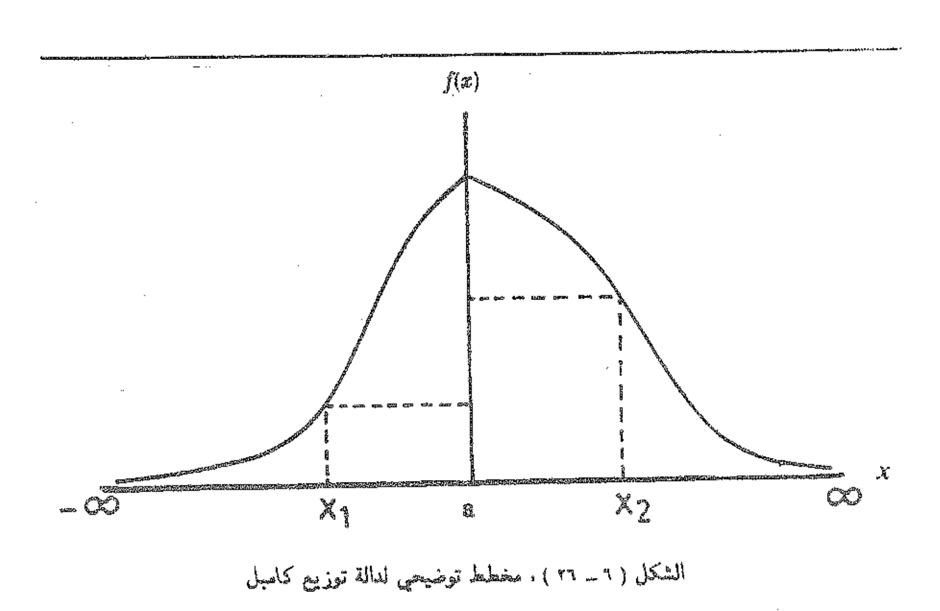
ويلاحظ ان هاتين النقطتين تقعان على بعد متساو الى يمين ويسار المنوال .

٤ ـ ان منحنى دالة توزيع كامبل ذا التواء موجب دائماً وذلك واضح من خلال مايلي :

$$S_k = \frac{\text{mean} - \text{mode}}{\sigma_x} = \frac{a + \lambda b - a}{\pi b / \sqrt{6}}$$

$$= \frac{\lambda \sqrt{6}}{\pi} \simeq 0{\cdot}45{\cdot}00534$$

ويلاحظ مما تقدم ان S_k (معامل الالتواء) مستقل عن b , a وذلك يعني انه مهما كانت قيمة كل من b , a فان شكل منحنى دالة هذا التوزيع هو ذاته من حيث مسار منحناه والمبين في الشكل (٦ ــ ٢٦) :



الشكل (٦ ــ ٢٦) : مخطط توضيحي لدالة توزيع كامبل

٦ ــ ٦ ــ ٨ : توزيع والْد Wald distribution

يقال ان المتغير العشوائي X يتوزع وفق توزيع والْد اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda (x-\mu)^2}{2\mu^2 x}\right\} ; x > 0$$

حيث ان λ , μ تمثلان معلمتي التوزيع بحيث ان $\mu, \lambda > 0$. علماً ان هنالك اشكال اخرى لهذا التوزيع لامجال لذكرها هنا والشكل الموضح اعلاه هو الشكل المعياري لتوزيع والد . ويعد هذا التوزيع واحداً من التوزيعات المهمة التي تدخل في موضوع التحليل المتسلسل **Sequential analysis** وفيما يلي بعض خصائص هذا التوزيع .

١ ـ ان الدالة التوزيعية لتوزيع والد هي :

$$F(x) = F^{*}\left\{(x-1)\sqrt{\frac{\lambda}{x\mu}}\right\} + e^{\frac{2\lambda}{\mu}} . F^{*}\left\{-(x+1)\sqrt{\frac{\lambda}{x\mu}}\right\}$$

حيث ان $F^{*}(\cdot)$ تمثل الدالة التوزيعية في التوزيع الطبيعي المعياري . وفي حالة ملاحظتنا ان x كبيرة نسبياً يمكن استخدام التقريب التالي :

$$F(x) \simeq 1 - e^{-\frac{\lambda}{2\mu}(x-2)} . \ln x$$

٢ ـ ان الدالة المولدة التراكمية هي :

$$K_X(t) = \frac{\lambda}{\mu}\left[1 - \left(1 - \frac{2\mu^2 t}{\lambda}\right)^{\frac{1}{2}}\right]$$

والتي من خلالها يمكن اثبات ان :

$$K_X'(0) = EX = \mu$$

وان

$$K_X''(0) = \sigma_X^2 = \frac{\mu^3}{\lambda}$$

وانه بشكل عام

$$K_X^{(r)}(0) = \prod_{j=2}^{r} (2j-3)\,\mu^{2r-1} . \lambda^{1-r} ; r \geq 2$$

تمارين عن التوزيعات المستمرة الاخرى

٦ ـ ٢٧ : اذا كان $X \sim \text{cauchy}(a, b)$ برهن ان

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x-a}{b}\right)$$

٦ ـ ٢٨ : اذا كان $X \sim \log N(\mu, \sigma^2)$. برهن ان

$$. Z = \frac{\ln X - \mu}{\sigma} \sim N(0, 1)$$

٦ ـ ٢٩ : افرض ان $K_X(t) = \ln M_X(t)$ تمثل الدالة المولدة التراكمية في التوزيع السوقي . بين ان

$$K'_X(0) = \alpha \quad , K''_X(0) = \frac{\pi^2 \beta^2}{3}$$

٦ ـ ٣٠ : في توزيع لابلاس وبفرض ان $\beta = 1, \alpha = 0$. برهن ان العزم المركزي ذو المرتبة r مساوٍ للصفر اذا كانت r عدد فردي . مساوٍ الى r! اذا كانت r عدد زوجي .

٦ ـ ٣١ : افرض ان X يتوزع وفق دالة توزيع وايبل بالمعلمتين a , b وان $Y = X^b$. برهن ان Y يتوزع كتوزيع اسي بالمعلمة a .

٦ ـ ٣٢ : اشتق صيغة العزم ذي المرتبة r حول نقطة الاصل في توزيع باريتو .

٦ ـ ٣٣ : برهن ان الدالة المولدة لعزوم توزيع كامبل حول نقطة الاصل هي $e^{at}\Gamma(1 - bt)$. (ملاحظة استفد من خصائص توزيع كاما) .

٦ ـ ٣٤ : بين ان الوسط والتباين في توزيع والدهما على التوالي $\mu^3 / \lambda, \mu$.

٦ ـ ٦ ـ ٩ : منظومة توزيعات پيرسون

Pearsonian system of distributions

افرض ان $f(x)$ تمثل دالة الكثافة الاحتمالية للمتغير العشوائي X ، وافرض ان $f(x)$ تحقق المعادلة التفاضلية التالية :

$$\frac{f'(x)}{f(x)} = \frac{x+a}{b_0 + b_1 x + b_2 x^2} \qquad \dots(*)$$

حيث a ، b_0 ، b_1 ، b_2 ثوابت حقيقية . عندئذٍ يقال ان $f(x)$ هي حالة خاصة من منظومة توزيعات پيرسون . وهنالك اشكال اخرى لهذه المنظومة لامجال لذكرها هنا ، وان العديد من التوزيعات التي سبق عرضها في الفقرات السابقة هي حالات خاصة من هذه المنظومة . فمثلاً يمكن اعتبار توزيع كاما بالمعلمتين α ، β حالة خاصة من منظومة توزيعات پيرسون كون ان هذا التوزيع يحقق $(*)$ وكما هو موضح بالآتي :

ان

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} . e^{-\frac{x}{\beta}}$$

فاذن

$$\ln f(x) = \ln\left(\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\right) + (\alpha-1)\ln x - \frac{x}{\beta}$$

وباشتقاق الطرفين نسبة الى x نحصل على :

$$\frac{f'(x)}{f(x)} = \frac{\alpha-1}{x} - \frac{1}{\beta} = \frac{\beta(\alpha-1)-x}{x\beta}$$

$$= \frac{x-\beta(\alpha-1)}{-x\beta} \qquad \dots(**)$$

وبمقارنة $(**)$ مع $(*)$ نجد ان

$$a = -\beta(\alpha-1) \quad , \quad b_0 = b_2 = 0 \quad , \quad b_1 = -\beta$$

وهذا يعني ان توزيع كاما هو حالة خاصة من منظومة توزيعات پيرسون .
كذلك يمكن اعتبار التوزيع الطبيعي حالة خاصة من هذه المنظومة وذلك لان :

$$\ln f(x) = \ln \left(\frac{1}{\sqrt{2\pi}\ \sigma} \right) - \frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2$$

وان

$$\frac{f'(x)}{f(x)} = -\left(\frac{x-\mu}{\sigma} \right)\left(\frac{1}{\sigma} \right) = \frac{x-\mu}{-\sigma^2} \quad \ldots (***)$$

وبمقارنة (***) مع (*) نجد ان

$$a = -\mu \quad , \quad b_0 = -\sigma^2 \quad , \quad b_1 = b_2 = 0$$

ويترك للقاريء البيان ان توزيع بيتا هو حالة خاصة من هذه المنظومة .

٦ ـ ٧ : التوزيعات المركبة Compound distributions

سبق وان تركزت دراستنا لموضوع التوزيعات الاحتمالية (سواء كانت متقطعة ام مستمرة) على توزيعات معلميه Parametric distributions يتم تشخيص كل واحد منها من خلال تحديد قيمة (او قيم) المعلمة (او المعالم) التي يتضمنها ذلك التوزيع . كذلك لاحظنا ان الدالة التوزيعية ، الدالة المولدة للعزوم ، العزوم وغيرها من المقاييس والخصائص الخاصة بتلك التوزيعات كانت تظهر بهيئة دوال تعتمد على بعض او جميع المعالم التي يتضمنها التوزيع وان تلك المعالم كانت تعتبر بحكم الثوابت في دالة التوزيع .

الا انه في الكثير من الاحوال (وخصوصاً في موضوع تقديرات بيز Bayes estimators في الاستدلال الاحصائي) نلاحظ ان معلمة أو معالم التوزيع تبدو هي الاخرى بحكم متغير عشوائي يسلك وفق دالة كتلة أو كثافة احتمالية . لذا وفي مثل هذه الاحوال يستوجب الامر استنتاج التوزيع الاحتمالي للمتغير العشوائي ، مثل X ، آخذين بنظر الاعتبار التوزيع الاحتمالي لمعلمة (او معالم) التوزيع .

ان التوزيع الجديد للمتغير X في هذه الحالة يسمى التوزيع المركب وان عملية استنتاج التوزيع الجديد تسمى خلط التوزيعات Mixture of distributions .

وعلى فرض ان $g(x;\theta)$ تمثل دالة الكتلة (او الكثافة) الاحتمالية للمتغير العشوائي X بالمعلمة θ وان $h(\theta)$ تمثل دالة الكتلة (او الكثافة) الاحتمالية الى θ معرفة على Ω_θ . عندئذٍ فان التوزيع المركب للمتغير X هو :

في حالة θ من النوع المتقطع
$$C(x) = \sum_{\Omega_\theta} g(x;\theta)h(\theta)$$

$$= \int_{\Omega_\theta} g(x;\theta)h(\theta)\,d\theta$$

وعلى نحو اكثر عمومية وبفرض ان $g(x;\theta_1,\theta_2,\ldots,\theta_k)$ تمثل دالة الكتلة (او الكثافة) الاحتمالية للمتغير X بالمعالم $\theta_1,\theta_2,\ldots,\theta_k$ وان $h(\theta_i)$ تمثل دالة الكتلة (او الكثافة) الاحتمالية الى θ_i معرفة على Ω_{θ_i} ، عندئذٍ فان التوزيع المركب الى X هو :

$$C(x;\theta_1,\theta_2,\ldots,\theta_{i-1},\theta_{i+1},\ldots\theta_k) = \sum_{\Omega_{\theta_i}} g(x;\theta_1,\theta_2,\ldots,\theta_k)h(\theta_i)$$ ،

عندما θ_i من النوع المتقطع . ويكون مساوياً الى :

$$\int_{\Omega_{\theta_i}} g(x;\theta_1,\theta_2,\ldots,\theta_k)h(\theta_i)\,d\theta_i$$ عندما θ_i من النوع المستمر .

وفيما يلي بعض الامثلة التوضيحية لاستنتاج التوزيعات المركبة :

٦ ـ ٧ ـ ١ : توزيع ثنائي الحدين المركب

Compound binomial distribution

افرض ان X متغير عشوائي يتوزع وفق دالة توزيع ثنائي الحدين بالمعلمتين n, P . عندئذٍ

$$P(x=r) = C_r^n P^r . q^{n-r} \quad ; r = 0, 1, 2, \ldots, n$$

الان بفرض ان n هي الاخرى متغير عشوائي يسلك وفق دالة توزيع پواسون بالمعلمة m . عندئذٍ :

$$h(n = k) = \frac{m^k . e^{-m}}{k!} \; ; K = 0, 1, 2, \ldots$$

وفي هذه الحالة يقال ان X يمتلك توزيع ثنائي الحدين المركب . ان دالة الكتلة الاحتمالية المشتركة الى X, n هي :

$$P(x = r, n = k) = h(n = k) . P(X = r \mid n = k)$$

$$= \frac{m^k . e^{-m}}{k!} . C_r^k P^r . q^{k-r}$$

وحيث ان $P(x = r \mid n = k)$ تعني احتمال الحصول على r محاولة ناجحة من بين k محاولة مستقلة فان ذلك يعني ان $k \geq r$. فاذن :

$$C(x = r) = \sum_{k=r}^{\infty} P(x = r, n = k)$$

$$= \sum_{k=r}^{\infty} \frac{m^k . e^{-m}}{K!} . C_r^k P^r . q^{k-r}$$

$$= e^{-m} . P^r \sum_{k=r}^{\infty} \frac{m^k}{K!} . \frac{K!}{r!(K-r)!} . q^{k-r}$$

$$= \frac{(mp)^r . e^{-m}}{r!} \sum_{k=r}^{\infty} \frac{(mq)^{k-r}}{(K-r)!}$$

$$= \frac{(mp)^r . e^{-m}}{r!} . \sum_{y=0}^{\infty} \frac{(mq)^y}{y!} , y = K - r$$

$$= \frac{(mp)^r . e^{-m}}{r!} . e^{mq} , q = 1 - p$$

$$\therefore C(x = r) = \frac{(mp)^r . e^{-mp}}{r!} \; ; r = 0, 1, 2, \ldots$$

والشكل الاخير يمثل دالة توزيع پواسون بالمعلمة (mp) .

٦ ـ ٧ ـ ٢ : توزيع ثنائي الحدين ـ بيتا المركب

Compound binomial-beta dist.

ليكن X متغير عشوائي يتوزع كتوزيع ثنائي الحدين بالمعلمتين n , p . عندئذٍ فان

$$P(x = r) = C_r^n P^r \cdot q^{n-r} \; ; r = 0, 1, 2, \ldots, n$$

وبفرض ان المعلمة P هي متغير عشوائي تتوزع وفق دالة توزيع بيتا بالمعلمتين α, β ، عندئذٍ

$$h(P) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} P^{\alpha-1} . (1-P)^{\beta-1} \; ; 0 < P < 1$$

فاذن

$$C(x = r; n, \alpha, \beta) = \int_0^1 P(x = r) . h(P) dP$$

$$= C_r^n \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} \int_0^1 P^r (1-P)^{n-r} . P^{\alpha-1} (1-P)^{\beta-1} dP$$

$$= C_r^n \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 P^{(\alpha+r)-1} . (1-P)^{(n+\beta-r)-1} dP$$

$$= C_r^n . \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} . \frac{\Gamma(\alpha+r).\Gamma(n+\beta-r)}{\Gamma(n+\alpha+\beta)} \quad ; r = 0, 1, 2, \ldots, n$$

والصيغة الاخيرة هي شكل من اشكال توزيع Po'lya – Eggenberger .

٦ ـ ٧ ـ ٣ : توزيع پواسون المركب

Compound poisson distribution

افرض ان X متغير عشوائي يتوزع وفق دالة توزيع پواسون بالمعلمة m عندئذٍ

$$P(x = r) = \frac{m^r . e^{-m}}{r!} \; ; r = 0, 1, 2, \ldots$$

وبفرض ان المعلمة m هي متغير عشوائي تسلك وفق دالة توزيع كاما بالمعلميتن α , β ، عندئذٍ :

$$h(m) = \frac{1}{\Gamma(\alpha).\beta^{\alpha}} . m^{\alpha-1} . e^{-\frac{m}{\beta}} \quad ; m > 0$$

فاذن

$$C(x = r ; \alpha, \beta) = \int_0^{\infty} P(x = r) . h(m) dm$$

$$= \frac{1}{r!\,\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} m^{(r+\alpha)-1} . e^{-\frac{m}{\beta'}} dm \; ; \beta' = \frac{\beta}{\beta+1}$$

$$= \frac{1}{r!\,\Gamma(\alpha)\beta^{\alpha}} . \Gamma(r+\alpha) . \left(\frac{\beta}{\beta+1} \right)^{r+\alpha}$$

$$= \frac{\Gamma(r+\alpha)}{r!\,\Gamma(\alpha)} . \frac{\beta^{r}}{(\beta+1)^{r+\alpha}}$$

$$= C_r^{r+\alpha-1} . \left(\frac{\beta}{\beta+1} \right)^{r} . \left(\frac{1}{\beta+1} \right)^{\alpha}$$

وبوضع $p = \frac{\beta}{\beta+1}$ فان $q = 1 - p = \frac{1}{\beta+1}$ عليه فان

$$C(x = r ; \alpha, \beta) = C_r^{r+\alpha-1} . p^{r} . q^{\alpha} \; ; r = 0, 1, 2, \ldots$$

والصيغة الاخيرة هي دالة توزيع ثنائي الحدين السالب بالمعلمتين (α , P).
ووفق هذا التصور يمكن استنتاج العديد من التوزيعات المركبة بالاسلوب الموضح في الامثلة السابقة . فمثلاً يمكن استنتاج التوزيعات المركبة التالية :

١ ـ توزيع پواسون ـ المنتظم المستمر بفرض ان المعلمة $\lambda \sim Cu(a, b)$
٢ ـ توزيع پواسون ـ ثنائي الحدين السالب بفرض ان المعلمة $\lambda \sim Nb(r, P)$
٣ ـ توزيع پواسون ـ اللوغارتمي الطبيعي بفرض ان المعلمة $\lambda \sim Log\,N(\mu, \sigma^2)$
٤ ـ توزيع پواسون ـ پواسون بفرض ان المعلمة $\lambda \sim Poisson(\theta)$
٥ ـ توزيع ثنائي الحدين السالب ـ توزيع كاما بفرض ان المعلمة $r \sim G(\alpha, \beta)$
٦ ـ توزيع طبيعي ـ طبيعي بفرض ان المعلمة $\mu \sim N(\theta_1, \theta_2)$

٧ـ توزيع طبيعي ـ كاما بفرض ان المعلمة $\sigma^2 \sim G(\alpha, \beta)$

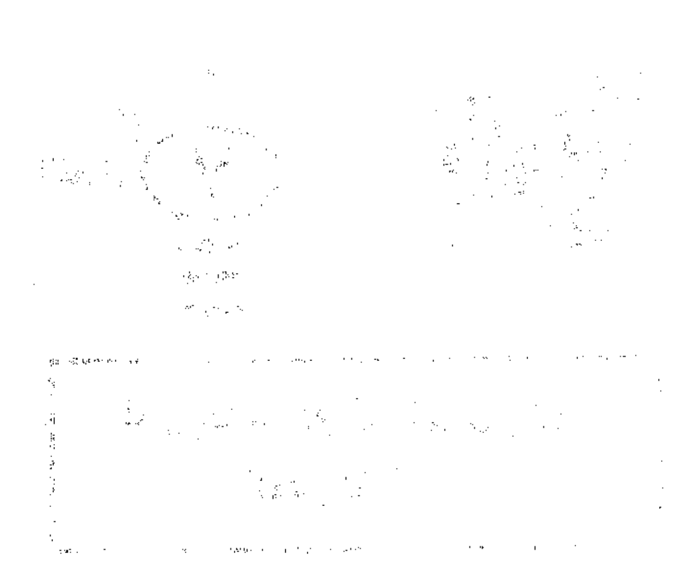
٨ ـ توزيع كاما ـ كاما بفرض ان المعلمة $\beta \sim G(\theta_1, \theta_2)$

ونترك مسألة استنتاج التوزيعات المركبة اعلاه كتمارين للقاريء .





توزيعات دوال المتغيرات العشوائية



الفصل السابع
توزيعات دوال المتغيرات العشوائية

Distributions of functions of random variables

لقد تركزت دراستنا في الفصلين الخامس والسادس على استعراض لاهم عوائل التوزيعات الاحتمالية النظرية الشائعة في النظرية الاحصائية مع بيان اهم خصائص هذه التوزيعات من حيث دوالها التوزيعية ، عزومها ، علاقة هذه التوزيعات ببعضها وغيرها من المقاييس ذات العلاقة بها .

في هذا الفصل سوف نركز الاهتمام على دراسة توزيعات دوال المتغيرات العشوائية من خلال عرض لاهم الاساليب المتاحة في استنتاج توزيعات هذه الدوال على الرغم من ان هذه الاساليب سبق وان استخدمت في بعض فقرات الفصلين الخامس والسادس .

٧ ــ ١ : توقعات دوال المتغيرات العشوائية :

افرض ان $X_1, X_2, ..., X_k$ متغيرات عشوائية وان $Y=g(X_1, X_2, ..., X_k)$ تمثل دالة بدلالة هذه المتغيرات . وافرض اننا نرغب في ايجاد توقع الدالة Y .

ان عملية ايجاد توقع الدالة Y هي في الحقيقة مكافئة الى $Eg(X_1, X_2, ..., X_k)$. وهذا يعني أن توقع Y يمكن التوصل اليه بطريقتين : الاولى هي استنتاج التوزيع الاحتمالي للمتغير Y ومن ثم حساب EY (اذا كان التوقع موجود) ، والثانية هي حساب توقع الدالة g على اساس التوزيع المشترك للمتغيرات $X_1, X_2, ..., X_k$. فعلى فرض ان المتغيرات المنوه عنها اعلاه من النوع المستمر ، فانه وعلى اساس تعريف التوقع الرياضي :

$$EY = \int_{-\infty}^{\infty} yf(y)\,dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_k).f(x_1, x_2, \ldots, x_k).dx_1.dx_2 \ldots dx_k$$

ومن الناحية التطبيقية علينا اختيار الطريقة الاسهل للتوصل الى توقع الدالة Y طالما ان النتيجة واحدة في كلا الحالتين . فمثلاً نرى انه من الافضل استنتاج توزيع المتغير Y ومن ثم حساب توقع هذا المتغير ، في حين لو تم اختيار الاسلوب الثاني فان ذلك يستوجب التعامل مع تكاملات عديدة الامر الذي قد يؤدي الى بعض الصعوبات في حساب توقع الدالة Y.

مثال (١) : افرض ان X_1, X_2 متغيران عشوائيان بدالة كثافة احتمالية مشتركة :

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

$-\infty < x_1, x_2 < \infty$. وافرض ان $Y = 2X_1 + 3X_2$ عندئذ فان

$$EY = E(2X_1 + 3X_2) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (2x_1 + 3x_2).$$
$$e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}.dx_1\,dx_2$$

$$= \frac{2}{\sqrt{2\pi}\,\sigma_1} \int_{-\infty}^{\infty} x_1.e^{\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} dx_1$$
$$+ \frac{3}{\sqrt{2\pi}\,\sigma_2} \int_{-\infty}^{\infty} x_2.e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2} dx_2 \quad \ldots(*)$$

ان التكامل الاول فى (*) ماهو الا ضعف توقع X_1 , $X_1 \sim N(\mu_1, \sigma_1^2)$
في حين ان التكامل الثاني في (*) ماهو الا ثلاثة امثال توقع توقع X_2 ،
$X_2 \sim N(\mu_2, \sigma_2^2)$. وهذا يعني ان

$$\begin{aligned} EY &= 2EX_1 + 3EX_2 \\ &= 2\mu_1 + 3\mu_2 \end{aligned}$$

لاحظ من هذا المثال ان هنالك بعض الصعوبة في التوصل الى توقع Y ، في حين وكما نعلم فان $f(x_1, x_2)$ تمثل دالة الكثافة الاحتمالية المشتركة لمتغيرين مستقلين هما $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$. وهذا يعني انه يمكن استنتاج توزيع المتغير Y ، وفق ما هو موضح في الفقرة (٦ ــ ٢ ــ ٦) ، على انه توزيع طبيعي بوسط قدره $(2\mu_1 + 3\mu_2)$ وتباين مقداره $(4\sigma_1^2 + 9\sigma_2^2)$ وهذا يعني ان $EY = 2\mu_1 + 3\mu_2$. لاحظ السهولة في عملية حساب توقع الدالة Y

٧ ــ ١ ــ ١ : الوسط والتباين لمجموع عدة متغيرات عشوائية .

سبق وان لاحظنا في فقرات عديدة من الفصلين الخامس والسادس عملية استنتاج عزوم مجموع عدة متغيرات عشوائية مستقلة تسلك وفق توزيع احتمالي معين (قد تكون بنفس معالم ذلك التوزيع او بمعالم مختلفة) ، على سبيل المثال لاحظ الاستنتاج في الفقرة (٦ ــ ٢ ــ ٦) . في هذه الفقرة سوف نستعرض عملية حساب الوسط والتباين لمجموع عدة متغيرات عشوائية وبشكل عام سواء كانت هذه المتغيرات مستقلة ام مرتبطة .

افرض ان $X_1, X_2, \ldots, X_k$ متغيرات عشوائية كل منها يتوزع وفق دالة كتلة (أو كثافة) احتمالية حدية مثل $P_i(x_i)$ (أو $f_i(x_i)$) .
عندئذٍ :

١ ــ ان $$E \sum_{i=1}^{k} X_i = \sum_{i=1}^{k} EX_i$$

البرهان :

$$E\sum_{i=1}^{k} X_i = \int_{x_1}\int_{x_2}\cdots\int_{x_k} (x_1 + x_2 + \ldots + x_k).f(x_1, x_2, \ldots, x_k)dx_1 dx_2 \ldots dx_k$$

$$= \int_{x_1}\int_{x_2}\cdots\int_{x_k} x_1 . f(x_1, x_2, \ldots, x_k)\, dx_1 dx_2 \ldots dx_k$$

$$+ \int_{x_1}\int_{x_2}\cdots\int_{x_k} x_2 . f(x_1, x_2, \ldots, x_k)\, dx_1 dx_2 \ldots dx_k + \ldots$$

$$+ \int_{x_1}\int_{x_2}\cdots\int_{x_k} x_k . f(x_1, x_2, \ldots, x_k)\, dx_1 dx_2 \ldots dx_k$$

لكن وبشكل عام فان :

$$\int_{x_1}\int_{x_2}\cdots\int_{x_i}\cdots\int_{x_k} x_i . f(x_1, x_2, \ldots, x_k)\, dx_1 . dx_2 . \ldots dx_i . \ldots . dx_k$$

$$= \int_{x_i} x_i . f_i(x_i)\, dx_i = EX_i$$

فاذن

$$E\sum_{i=1}^{k} X_i = \sum_{i=1}^{k} EX_i$$

ونفس خطوات البرهان تتم لحالة المتغيرات المتقطعة بمجرد استبدال رمز التكامل برمز الجمع .

مثال (٢) : اذا علمت ان $X_1 \sim N(2, 6)$, $X_2 \sim N(3, 8)$, $X_3 \sim N(6, 10)$. وافرض ان $Y = X_1 + X_2 + X_3$. جد EY .

الحل :

$$EY = E(X_1 + X_2 + X_3) = EX_1 + EX_2 + EX_3$$
$$= 2 + 3 + 6$$
$$= 11$$

مثال (٣) : افرض ان $X_1 \sim b\left(6, \frac{1}{2}\right)$ ، $X_2 \sim b\left(9, \frac{1}{3}\right)$ ،

$X_3 \sim b\left(12, \frac{1}{14}\right)$. وافرض ان $Y = X_1 + X_2 + X_3$. جد EY .

الحل :

$$EY = EX_1 + EX_2 + EX_3$$
$$= n_1P_1 + n_2P_2 + n_3P_3$$
$$= 6\left(\frac{1}{2}\right) + 9\left(\frac{1}{3}\right) + 12\left(\frac{1}{4}\right)$$

$$\therefore EY = 3 + 3 + 3 = 9$$

٢ ـ ان

$$V\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k} \sigma_i^2 + 2\sum_{i}\sum_{<j} \sigma_{ij}$$

حيث ان σ_{ij} تعني التباين المشترك بين المتغيرين X_j, X_i

البرهان :

$$V\left(\sum_{i=1}^{k} X_i\right) = E\left[\left(\sum_{i=1}^{k} X_i\right) - E\left(\sum_{i=1}^{k} X_i\right)\right]^2$$

$$= E\left[\sum_{i=1}^{k} X_i - \sum_{i=1}^{k} EX_i\right]^2$$

$$= E\left[\sum_{i=1}^{k}(X_i - EX_i)\right]^2$$

$$= E\left[\sum_{i=1}^{k}(X_i - EX_i)^2 + 2\sum\sum_{i<j}(X_i - EX_i)(X_j - EX_j)\right]$$

$$= \sum_{i=1}^{k}E(X_i - EX_i)^2 + 2\sum\sum_{i<j}E(X_i - EX_i)(X_j - EX_j)$$

$$= \sum_{i=1}^{k}\sigma_i^2 + 2\sum\sum_{i<j}\sigma_{ij}$$

أو

$$= \sum_{i=1}^{k}\sigma_i^2 + 2\sum\sum_{i<j}\sigma_i \cdot \sigma_j \cdot \rho_{ij}$$

حيث ρ_{ij} تعني معامل الارتباط بين المتغيرين X_j , X_i
ويتضح مما تقدم انه اذا كانت هذه المتغيرات مستقلة تصادفياً فذلك يعني ان $\sigma_{ij} = 0$. وعندئذٍ فان

$$V\left(\sum_{i=1}^{k}X_i\right) = \sum_{i=1}^{k}V(X_i) = \sum_{i=1}^{k}\sigma_i^2$$

مثال (٤) : اذا علمت ان $X_1 \sim N(2,4)$, $X_2 \sim N(3,9)$,
وان $X_3 \sim N(5,16)$, $\rho_{12} = 0{\cdot}4$, $\rho_{13} = 0{\cdot}6$, $\rho_{23} = 0{\cdot}3$.
جد تباين $Y = X_1 + X_2 + X_3$.

الحل :

$$V(Y) = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + 2(\sigma_1\sigma_2\rho_{12} + \sigma_1\sigma_3\rho_{13} + \sigma_2\sigma_3\rho_{23})$$
$$= 4 + 9 + 16 + 2[(2)(3)(0{\cdot}4) + (2)(4)(0{\cdot}6) + (3)(4)(0{\cdot}3)]$$
$$= 50{\cdot}6$$

مثال (٥) : لمعطيات المثال (٤) وبفرض ان هذه المتغيرات مستقلة تصادفياً ، عندئذٍ فان .

$$V(Y) = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 = 29$$

وبشكل عام اذا كانت $a_1, a_2, \ldots, a_k$ ثوابت حقيقية فان

$$V\left(\sum_{i=1}^{k} a_i X_i\right) = \sum_{i=1}^{k} a_i^2 \sigma_i^2 + 2 \sum_{i<j}\sum a_i a_j \sigma_{ij}$$

وبشكل خاص يمكن الاستنتاج بان :

$$V(X_1 \pm X_2) = \sigma_1^2 + \sigma_2^2 \pm \sigma_{12}$$

٣ ـ اذا كانت $a_i, b_j ; i = 1, 2, \ldots, k$
$j = 1, 2, \ldots, m$ ثوابت حقيقية فان

$$Cov\left(\sum_{i=1}^{k} a_i X_i , \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j \sigma_{ij}$$

البرهان : ليكن μ_i يمثل الوسط الى X_i وان μ_j^* يمثل الوسط الى Y_j . عندئذٍ

$$Cov\left(\sum_{i=1}^{k} a_i X_i , \sum_{j=1}^{m} b_j Y_j\right) = E\left[\sum_{i=1}^{k} a_i X_i - \sum_{i=1}^{k} a_i \mu_i\right]\left[\sum_{j=1}^{m} b_j Y_j - \sum_{j=1}^{m} b_j \mu_j^*\right]$$

$$= E\left[\sum_{i=1}^{k} a_i (X_i - \mu_i)\right]\left[\sum_{j=1}^{m} b_j (Y_j - \mu_j^*\right]$$

$$= E \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_i (X_i - \mu_i)(Y_j - \mu_j^*)$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j E(X_i - \mu_i)(Y_j - \mu_j^*)$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j \cdot \sigma_{ij}$$

او

$$= \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j \cdot \sigma_i \sigma_j \rho_{ij}$$

حيث ان ρ_{ij} تعني معامل الارتباط بين المتغيرين Y_j, X_i .

مثال (٩) : اذا علمت ان $X_1 \sim N(1,4)$ $X_2 \sim N(2,9)$

$Y_1 \sim N(3,4)$, $Y_2 \sim N(5,16)$

وان $\rho_{x_1y_1} = 0.3$, $\rho_{x_1y_2} = 0.4$, $\rho_{x_2y_1} = 0.6$; $\rho_{x_2y_2} = 0.5$

وان $Z_1 = 2X_1 + 3X_2$, $Z_2 = 4Y_1 + 3Y_2$ جد $Cov(Z_1, Z_2)$

الحل :

$$\begin{aligned} Cov(Z_1, Z_2) &= a_1 b_1 \sigma_{x_1} \sigma_{y_1} \rho_{x_1y_1} + a_1 b_2 \sigma_{x_1} \sigma_{y_2} \sigma_{x_1y_2} \\ &+ a_2 b_1 \sigma_{x_2} \sigma_{y_1} \rho_{x_2y_1} + a_2 b_2 \sigma_{x_2} \sigma_{y_2} \rho_{x_2y_2} \\ &= (2)(4)(2)(2)(0.3) + (2)(3)(2)(4)(0.4) \\ &+ (3)(4)(3)(2)(0.6) + (3)(3)(3)(4)(0.5) \\ \therefore Cov(Z_1, Z_2) &= 126 \end{aligned}$$

٤ ـ اذا كانت $X_1, X_2, \ldots, X_k$ متغيرات عشوائية مستقلة تتوزع وفق نفس دالة الكثافة الاحتمالية $f(x)$ او دالة كتلة احتمالية $P(x)$ ، وان $\bar{X}$ يمثل الوسط الحسابي لهذه المتغيرات وان $EX_i = \mu, \sigma_i^2 = \sigma^2 \forall i = 1, 2, \ldots, k$

عندئذ :

$$E\bar{X} = \mu, V(\bar{X}) = \frac{\sigma^2}{K}$$

البرهان :

$$E\bar{X} = \frac{1}{K} E \sum_{i=1}^{k} X_i = \frac{1}{K} \sum_{i=1}^{k} EX_i = \frac{1}{K} \sum_{i=1}^{k} \mu = \mu$$

كذلك فان

$$V(\bar{X}) = E(\bar{X} - E\bar{X})^2 = E(\bar{X} - \mu)^2$$

$$= \frac{1}{K^2} E\left[\sum_{i=1}^{k} X_i - K\mu \right]^2$$

$$= \frac{1}{K^2} E\left[\sum_{i=1}^{k} (X_i - \mu) \right]^2$$

$$= \frac{1}{K^2} E\left[\sum_{i=1}^{k} (X_i - \mu)^2 + 2 \sum\sum_{i<j} (X_i - \mu)(X_j - \mu) \right]$$

$$= \frac{1}{K^2} \sum_{i=1}^{k} E(X_i - \mu)^2 , E(X_i - \mu)(X_j - \mu) = 0$$ (بالفرض)

$$= \frac{1}{K^2} \sum_{i=1}^{k} \sigma^2 = \frac{\sigma^2}{K}$$

٧ ــ ١ ــ ٢ : الوسط والتباين لحاصل ضرب أو قسمة متغيرين

افرض ان Y , X متغيران بدالة كثافة احتمالية مشتركة $f(x, y)$ أو كتلة احتمالية مشتركة $P(x, y)$. وعلى فرض ان العزوم الحدية والمشتركة موجودة عندئذ ،

١ ــ ان

$$EX . Y = \mu_x . \mu_y + \sigma_{xy}$$

البرهان :

$$\begin{aligned} EXY &= EX.Y - \mu_x \cdot \mu_y + \mu_x \cdot \mu_y \\ &= E(X - \mu_x)(Y - \mu_y) + \mu_x \cdot \mu_y \\ &= \sigma_{xy} + \mu_x \cdot \mu_y \end{aligned}$$

أو

$$= \sigma_x \cdot \sigma_y \cdot \rho_{xy} + \mu_x \cdot \mu_y$$

مثال (٧) : اذا علمت ان X, Y متغيران عشوائيان وان $\mu_x = 2$

$\rho_{xy} = -0{\cdot}6, \sigma_y = 6, \sigma_x = 4, \mu_y = 3$ جد EXY

الحل :

$$\begin{aligned} EXY &= \sigma_x \cdot \sigma_y \cdot \rho_{xy} + \mu_x \cdot \mu_y \\ &= (4)(6)(-0{\cdot}6) + (2)(3) \\ &= -8{\cdot}4 \end{aligned}$$

٢ ــ ان

$$V(XY) = \mu_y^2 \sigma_x^2 + \mu_x^2 \sigma_y^2 + 2\mu_x \mu_y \sigma_{xy} - \sigma_{xy}^2 + G$$

البرهان $V(XY) = E(XY)^2 - (EXY)^2$

لاغراض السهولة في الاشتقاق فاننا سوف نكتب $X.Y$ بالشكل التالي :

$$X.Y = \mu_x \mu_y + (X - \mu_x)\mu_y + (Y - \mu_y)\mu_x + (X - \mu_x)(Y - \mu_y)$$

فاذن

$$V(XY) = E[\mu_x \mu_y + (X - \mu_x)\mu_y + (Y - \mu_y)\mu_x + (X - \mu_x)(Y - \mu_y)]^2 - (EXY)^2$$

وبفتح القوس الكبير والتعويض عن EXY بما يساويه نحصل على :

$$V(XY) = \mu_y^2 \sigma_x^2 + \mu_x^2 \sigma_y^2 + 2\mu_x \mu_y \sigma_{xy} - \sigma_{xy}^2 + G$$

حيث ان

$$G = E(X - \mu_x)^2 (Y - \mu_y)^2 + 2\mu_y E(X - \mu_x)^2 (Y - \mu_y)$$
$$+ 2\mu_x E(X - \mu_x)(Y - \mu_y)^2$$

وبشكل خاص اذا كان Y , X مستقلين عندئذ :

$$EXY = EX . EY = \mu_x \cdot \mu_y$$

وان

$$V(XY) = \mu_y^2 \sigma_x^2 + \mu_x^2 \sigma_y^2 + \sigma_x^2 \cdot \sigma_y^2$$

مثال (٨) : اذا علمت ان $X \sim N(1,4)$ مستقل عن $Y \sim N(2,3)$. جد توقع حاصل ضرب X في Y وكذلك تباين حاصل الضرب .

الحل :

$$EXY = EX . EY = (1)(2) = 2$$

$$V(XY) = \mu_y^2 \cdot \sigma_x^2 + \mu_x^2 \cdot \sigma_y^2 + \sigma_x^2 \cdot \sigma_y^2$$

$$= (2^2)(4) + (1^2)(3) + (4)(3) = 31$$

٣ ـ من الصعوبة تحديد صيغة للقيمة المتوقعة والتباين لحاصل قسمة متغيرين في حالة كونهما مرتبطين . الا انه امكن ايجاد صيغة تقريبية لكل منهما وهي :

$$E\left(\frac{X}{Y}\right) \simeq \frac{\mu_x}{\mu_y} + \frac{\mu_x}{\mu_y^3} \cdot \sigma_y^2 - \frac{1}{\mu_y^2} \cdot \sigma_{xy}$$

وان

$$V\left(\frac{X}{Y}\right) \simeq \left(\frac{\mu_x}{\mu_y}\right)^2 \cdot \left[\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2} - \frac{2\sigma_{xy}}{\mu_x \cdot \mu_y}\right]$$

مثال (٩) : اذا علمت ان Y,X متغيران عشوائيان بدالة كثافة إحتمالية مشتركة معينة ، وانه توفرت لديك المعلومات التالية عن هذا التوزيع :

$\mu_x = 2, \mu_y = 3, \sigma_x^2 = 4, \sigma_y^2 = 5, \sigma_{xy} = 4$

جد V(X/Y), E(X/Y) .

الحل :

$$E\left(\frac{X}{Y}\right) \simeq \frac{2}{3} + \frac{2}{27} . 5 - \frac{1}{9} . 4 = \frac{16}{27}$$

$$V\left(\frac{X}{Y}\right) \simeq \left(\frac{2}{3}\right)^2 . \left[\frac{4}{4} + \frac{5}{9} - \frac{2(4)}{(2)(3)}\right]$$

$$= \frac{8}{81}$$

وبشكل خاص اذا كان Y,X مستقلين فان :

$$E\left(\frac{X}{Y}\right) = EX . E\frac{1}{Y} = \frac{\mu_x}{H_y}$$

حيث ان H_y تعني الوسط التوافقي الى Y . وان

$$V\left(\frac{X}{Y}\right) = E\left(\frac{X}{Y}\right)^2 - \frac{\mu_x^2}{H_y^2} = EX^2 . E\frac{1}{Y^2} - \frac{\mu_x^2}{H_y^2}$$

$$= (\sigma_x^2 + \mu_x^2) E\frac{1}{Y^2} - \frac{\mu_x^2}{H_y^2}$$

$$= \sigma_x^2 E\frac{1}{Y^2} + \mu_x^2 E\left[\left(\frac{1}{Y}\right)^2 - \left(\frac{1}{H_y}\right)^2\right]$$

لكن

$$E\left(\frac{1}{Y}\right)^2 - E\left(\frac{1}{H_y}\right)^2 = E\left(\frac{1}{Y}\right)^2 - \left[E\left(\frac{1}{Y}\right)\right]^2$$

$$= \sigma^2_{1/Y}$$

فاذن

$$E\left(\frac{1}{Y^2}\right) = \sigma^2_{1/Y} + \frac{1}{H_y^2}$$

عليه فان

$$V\left(\frac{X}{Y}\right) = \sigma_x^2\left(\sigma^2_{1/Y} + \frac{1}{H_y^2}\right) + \mu_x^2\,\sigma^2_{1/Y}$$

$$= \frac{\sigma^2{}_x}{H_y^2} + \sigma^2_{1/Y}(\sigma_x^2 + \mu_x^2)$$

مثال (١٠) : افرض ان $X \sim G(2,3)$ مستقل عن $Y \sim G(3,5)$ جد

$$V\left(\frac{X}{Y}\right), E\left(\frac{X}{Y}\right)$$

الحل :

حيث ان $X \sim G(2,3)$ فذلك يعني ان $\mu_x = \alpha.\beta = 6$ وان $\sigma_x^2 = \alpha.\beta^2 = 18$.

كذلك فان $Y \sim G(3,5)$ فذلك يعني ان :

$$f(y) = \frac{1}{\Gamma(3).5^3}\, y^2 . e^{-\frac{y}{5}} = \frac{1}{250}\, y^2 . e^{-\frac{y}{5}} \;;\; y > 0$$

فاذن

$$\frac{1}{H_y} = E\left(\frac{1}{Y}\right) = \frac{1}{250}\int_0^\infty y . e^{-\frac{y}{5}}\, dy$$

$$= \frac{1}{250} \cdot \Gamma(2) \cdot 5^2 = \frac{1}{10} \qquad \therefore H_y = 10$$

كذلك فان :

$$E\left(\frac{1}{Y^2}\right) = \frac{1}{250} \int_0^{\infty} e^{-\frac{y}{5}} \, dy = \frac{1}{50}$$

عليه فان :

$$\sigma^2_{1/y} = E\left(\frac{1}{Y^2}\right) - \left[E\left(\frac{1}{Y}\right)\right]^2$$

$$= \frac{1}{50} - \frac{1}{100} = \frac{1}{100}$$

وبذلك فان :

$$E\left(\frac{X}{Y}\right) = \frac{\mu_x}{H_y} = \frac{6}{10} = 0.6$$

وان

$$V\left(\frac{X}{Y}\right) = \frac{\sigma_x^2}{H_y^2} + \sigma^2_{1/y}(\sigma_x^2 + \mu_x^2)$$

$$= \frac{18}{100} + \frac{1}{100}(18 + 36) = \frac{18}{25} = 0.72$$

٧ ـ ٢ ـ استنتاج التوزيعات باستخدام الدالة التوزيعية

سبقت الاشارة لموضوع استنتاج التوزيعات الاحتمالية باستخدام الدالة التوزيعية في الفقرة (١ ـ ٥ ـ ٢) وافترضنا في حينه ان يكون المتغير العشوائي من النوع المستمر . وفيما يلي وصف لذلك .

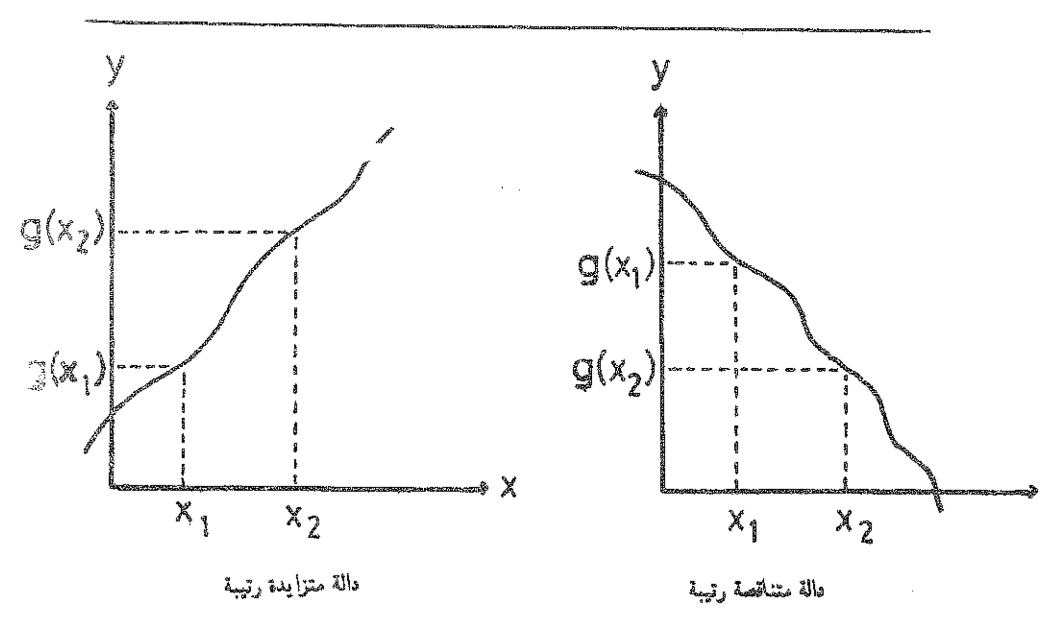
افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x)$ ودالة توزيعية $F(x)$. وافرض ان $Y = g(X)$ تمثل دالة بدلالة X بحيث ان $g(X)$ دالة وحيدة القيمة ، واننا نرغب في استنتاج التوزيع الاحتمالي للمتغير Y ، اي ايجاد دالة الكثافة الاحتمالية للمتغير Y ولتكن $f(y)$. بشكل عام فان الدوال التي سنتعامل معها في هذه الفقرة تقسم الى نوعين رئيسين هما :

١ ـ دوال متزايدة رتيبة Monotonically increasing functions

يقال للدالة $Y = g(X)$ انها دالة مستمرة متزايدة رتيبة اذا كانت $g(x_1) > g(x_2)$ لكل $x_1 > x_2$ ، حيث x_2 , x_1 قيم معطاة الى x . على سبيل المثال فان الدالة $Y = X^2$ هي دالة مستمرة ومتزايدة رتيبة في الفترة $(0 , \infty)$.

٢ ـ دوال متناقصة رتيبة Monotonically decreasing functions

يقال للدالة $Y = g(X)$ انها دالة مستمرة متناقصة رتيبة اذا كانت $g(x_1) < g(x_2)$ لكل $x_1 > x_2$. على سبيل المثال فان الدالة $Y = \frac{1}{X^2}$ هي دالة متناقصة رتيبة في الفترة $(1 , \infty)$. واذا كانت $g(x_1) = g(x_2)$ لكل $x_1 > x_2$ او $x_1 < x_2$ عندئذٍ يقال للدالة $Y = g(X)$ انها دالة غير متناقصة رتيبة او دالة غير متزايدة رتيبة . والشكل (٧ ـ ١) يوضح هذا النوع من الدوال .



الشكل (٧ ـ ١) : توضيح للدوال المتناقصة والمتزايدة الرتيبة

الان بفرض ان الدالة $Y = g(X)$ متزايدة رتيبة وان معكوس هذه الدالة هو $X = g^{-1}(Y)$. ان كل نقطة معرفة على المحور X تعرف نقطة واحدة فقط على المحور y من خلال $Y = g(x)$ وان كل نقطة من المحور y تعرف نقطة واحدة فقط على المحور X من خلال $X = g^{-1}(Y)$ ، وهذا يعني ان الحادثة المقابلة للحادثة $\{Y \leq y_0\}$ في فضاء المتغير X هي $\{X \leq x_0\}$ ، حيث ان $y_0 = g(x_0)$ ، والعكس صحيح ايضاً . ووفق هذا المفهوم يمكن استنتاج توزيع المتغير Y اذا علم توزيع المتغير X . ان الخطوات الرئيسية في ايجاد توزيع Y هي :

١ ـ يتم صياغة الحادثة $\{Y \leq y\}$ بدلالة الحادثة المقابلة لها في فضاء X .

٢ ـ ايجاد الدالة التوزيعية الى y اي $F(y)$.

٣ ـ ايجاد مشتقة الدالة $F(y)$ نسبة الى y . هذه المشتقة ماهي الا دالة الكثافة الاحتمالية للمتغير Y استناداً الى العلاقة ما بين دوال الكثافة الاحتمالية والدوال التوزيعية المنوه عنها في الفقرة (١ ـ ٥ ـ ٢) .

٤ ـ تحديد فضاء المتغير Y على ضوء الدالة $Y = g(X)$.

وبالرموز يمكن التعبير عن هذه الخطوات بما يلي :

$$F_Y(y) = P_r(Y \leq y) = P_r(Y \leq g(x))$$
$$= P_r(X \leq g^{-1}(y))$$
$$= F_X(g^{-1}(y))$$

فاذن

$$f_Y(y) = \frac{dF_X(g^{-1}(y))}{dy}$$

وباستخدام قاعدة السلسلة في حساب التفاضل نحصل على

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{dg^{-1}(y)}{dy}, x = g^{-1}(y)$$

فاذن

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{dx}{dy}$$

اما اذا كانت الدالة $Y = g(X)$ متناقصة رتيبة عندئذٍ فان الحادثة المقابلة للحادثة $\{Y \leq y\}$ في فضاء X هي $\{X \geq g^{-1}(y)\}$. وهذا يعني ان :

$$F_Y(y) = P_r(Y \leq y) = P_r(X \leq g^{-1}(y))$$

$$= 1 - P_r(X \leq g^{-1}(y))$$

$$= 1 - F_X(g^{-1}(y))$$

وعندئذٍ فان

$$f_Y(y) = -f_X(x) \frac{dx}{dy}, x = g^{-1}(y)$$

وحيث ان دوال الكثافة الاحتمالية هي دوال غير سالبة فان ذلك يستوجب اهمال الاشارة السالبة التي تظهر بعد عملية الاشتقاق . لذلك وبشكل عام ولاية دالة مستمرة رتيبة سواء كانت متزايدة ام متناقصة فان

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left| \frac{dx}{dy} \right| \qquad \ldots(*)$$

حيث ان $\left| \frac{dx}{dy} \right|$ يسمى « معامل التحويل لجاكوبيان Jacobian » او بشكل مختصر « معامل تحويل الدالة $Y = g(X)$ » الذي سيرد ذكره لاحقاً .

مثال (١١) : افرض ان $f(x) = \frac{1}{x^2}; x \geq 1$ جد دالة الكثافة الاحتمالية للمتغير $Y = e^{-X}$

الحل :

$$F(y) = P_r(Y \leq y) = P_r(e^{-x} \leq y)$$

$$= P_r(-X \leq \ln y) = P_r(X \geq -\ln y)$$

$$= 1 - P_r(X \leq -\ln y) = 1 - F_X(-\ln y)$$

$$= 1 - \int_{1}^{-\ln y} \frac{1}{x^2} dx = 1 + \left[\frac{1}{x} \right]_{1}^{-\ln y}$$

$$\therefore F(y) = -(\ln y)^{-1}$$

$$\therefore f(y) = F'(y) = \frac{1}{y(\ln y)^2} \; ; 0 < y \leq e^{-1}$$

ويمكن التوصل لنفس الاستنتاج باستخدام الصيغة (*)

مثال (١٢) : افرض ان $X \sim N(0.1)$. جد دالة الكثافة الاحتمالية الى $Y = X^2$.

الحل :

ان

$$f_Y(y) = f_X(g^{-1}(y)) . \left| \frac{dx}{dy} \right|$$

وان

$$y = x^2 \rightarrow x = \pm \sqrt{y}$$

$$\frac{dx}{dy} = \pm \frac{1}{2\sqrt{y}} \rightarrow \left| \frac{dx}{dy} \right| = \frac{1}{2\sqrt{y}}$$

$$\therefore f_Y(y) = f_X(\sqrt{y}) . \frac{1}{2\sqrt{y}} + f_X(-\sqrt{y}) . \frac{1}{2\sqrt{y}}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} . \frac{1}{2\sqrt{y}}$$

$$= \frac{1}{\sqrt{2\pi}} . y^{-\frac{1}{2}} . e^{-\frac{1}{2}y} = \frac{1}{\Gamma\left(\frac{1}{2}\right)\sqrt{2}} y^{-\frac{1}{2}} e^{-\frac{1}{2}y}$$

$$, y > 0$$

والدالة الاخيرة ماهي الا دالة توزيع كاما المعرف بالمعلمتين $\beta = 2, \alpha = \frac{1}{2}$ ونترك للقاريء استخدام الخطوات الاربعة المنوة عنها سابقاً في التوصل الى $Y = X^2$.

مثال (١٣): افرض ان X متغير عشوائي يتوزع وفق دالة التوزيع المنتظم المستمر على الفترة $(0,1)$. جد دالة الكثافة الاحتمالية الى $Y = -2 \ln X$.

الحل :

$$F(Y) = P_r(Y \leq y) = P_r(-2 \ln X \leq y)$$

$$= P_r\left(\ln X \geq -\frac{1}{2} y\right) = P_r\left(X \geq e^{-\frac{1}{2} y}\right)$$

$$= 1 - P_r(X \leq e^{-\frac{1}{2} y}) = 1 - \int_0^{e^{-\frac{1}{2} y}} dx$$

$$= 1 - [x]_0^{e^{-\frac{1}{2} y}} \qquad = 1 - e^{-\frac{1}{2} y}$$

فاذن

$$f(y) = F'(y) = \frac{1}{2} e^{-\frac{1}{2} y} \quad ; y > 0$$

والدالة الاخيرة هي دالة توزيع اسي فيه $\theta = \frac{1}{2}$.

٧ – ٢ – ١ : توزيع حاصل جمع (او الفرق بين) متغيرين

افرض ان X, Y متغيران عشوائيان بدالة كثافة احتمالية مشتركة $f(x, y)$ وافرض ان $V = X - Y, Z = X + Y$. عندئذ

$$f(z) = \int_{\Omega_x} f(x, Z - x)\, dx \quad \ldots (1)$$

$$= \int_{\Omega_y} f(Z - y, y)\, dy \quad \ldots (2)$$

$$f(v) = \int_{\Omega_x} f(x, x - v)\, dx \quad \ldots (3)$$

$$= \int_{\Omega_y} f(v + y, y)\, dy \quad \ldots (4)$$

البرهان : سوف نبرهن الصيغة (١) فقط ونترك برهنة الصيغ الثلاث الباقية للقاريء .

ان

$$F(z) = P_r(Z \leq z) = P_r(X + Y \leq z)$$

$$= \iint_{x + y \leq z} f(x, y)\, dxdy$$

$$= \int_{\Omega_x} \left[\int_{-\infty}^{z - x} f(x, y)\, dy \right] dx$$

وبوضع $Y = Z - X$ عندئذٍ

$$F(z) = \int_{\Omega_x} \left[\int_{-\infty}^{z} f(x, Z-x)\, dz \right] dx$$

$$= \int_{-\infty}^{z} \left[\int_{\Omega_x} f(x, Z-x)\, dx \right] dz$$

وباشتقاق الطرفين نسبة الى Z نحصل على

$$f(z) = F'(z) = \int_{\Omega_x} f(x, Z-x)\, dx$$

وفي حالة كون المتغيرين X , Y مستقلين عندئذٍ :

$$f(z) = \int_{\Omega_x} f(z-x) . f(x)\, dx = \int_{\Omega_y} f(z-y) . f(y)\, dy \quad \ldots (5)$$

ويطلب من القاريء برهنة ذلك. ان الصيغة (5) غالباً ماتسمى « صيغة الالتفافية Convolution formula » او ماتسمى في بعض الاحيان التفافية الدالتين $f(y), f(x)$.

مثال (١٤) : افرض ان $X \sim N(\mu_1, \sigma_1^2)$ مستقل عن $Y \sim N(\mu_2, \sigma_2^2)$. جد التوزيع الاحتمالي الى $Z = X + Y$

الحل :

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma_1} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}, f(z-x) = \frac{1}{\sqrt{2\pi}\,\sigma_2} e^{-\frac{1}{2}\left(\frac{z-x-\mu_2}{\sigma_2}\right)^2}$$

$$\therefore f(z)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\,\sigma_1}e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}\cdot\frac{1}{\sqrt{2\pi}\,\sigma_2}e^{-\frac{1}{2}\left(\frac{z-x-\mu_2}{\sigma_2}\right)^2}$$

$$=\frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2+\left(\frac{z-x-\mu_2}{\sigma_2}\right)^2\right]}dx$$

الان بفتح القوسين الصغيرين داخل القوس الكبير وتجميع الحدود التي تتضمن x^2 وتلك التي تتضمن x نحصل على :

$$f(z)=\frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty}e^{-\frac{1}{2}Q}dx$$

حيث ان

$$Q=x_1^2\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)-2x\left(\frac{\mu_1}{\sigma_1^2}+\frac{z-\mu_2}{\sigma_2^2}\right)+\frac{(z-\mu_2)^2}{\sigma_2^2}+\frac{\mu_1^2}{\sigma_1^2}\ldots(*)$$

الان باكمال المربع في (*) بدلالة x واجراء التكامل نسبة لهذا المتغير نحصل على :

$$f(z)=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}e^{-\frac{[z-(\mu_1+\mu_2)]^2}{2(\sigma_1^2+\sigma_2^2)}}\quad,-\infty<z<\infty$$

ويلاحظ من الدالة الاخيرة ان $Z=X+Y$ متغير عشوائي يتوزع كتوزيع طبيعي بوسط قدره $(\mu_1+\mu_2)$ وتباين مقداره $(\sigma_1^2+\sigma_2^2)$.

٧ ـ ٢ ـ ٢ : توزيع حاصل ضرب وقسمة متغيرين .

افرض ان X, Y متغيران عشوائيان بدالة كثافة احتمالية مشتركة $f(x, y)$ وافرض $Z = X . Y$ ، $V = \dfrac{X}{Y}$ عندئذٍ .

$$f(z) = \int_{\Omega_x} |x|^{-1} . f\left(x, \frac{Z}{x}\right) dx$$

$$= \int_{\Omega_y} |y|^{-1} . f\left(\frac{Z}{y}, y\right) dy$$

وان

$$f(v) = \int_{\Omega_y} |y| . f(vy, y)\, dy$$

البرهان :

$$F(z) = P_r(Z \leq z) = P_r(X . Y \leq z)$$

$$= \iint_{x \cdot y \leq z} f(x, y)\, dx . dy$$

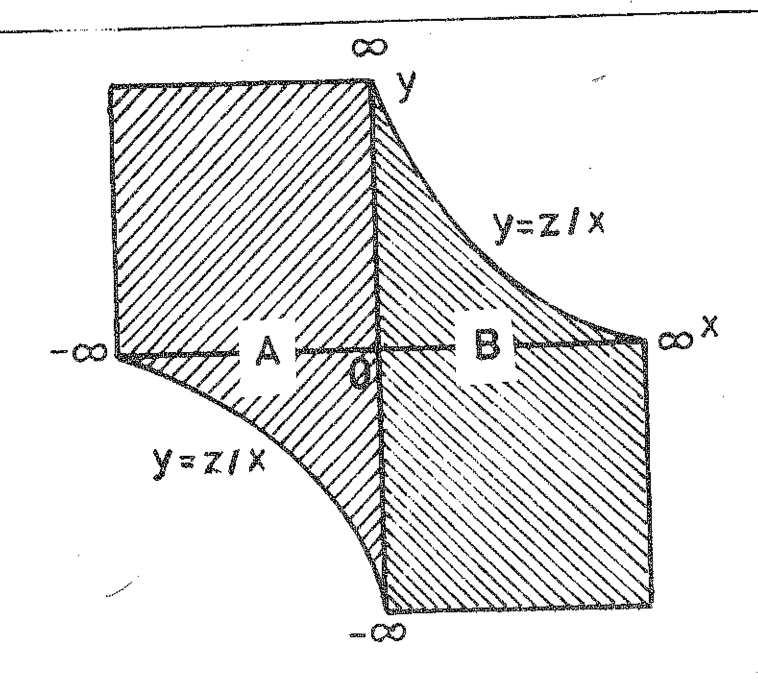
$$= \int_{-\infty}^{0} \left[\int_{z/x}^{\infty} f(x, y)\, dy \right] dx + \int_{0}^{\infty} \left[\int_{-\infty}^{z/x} f(x, y) \cdot dy \right] dx$$

لاحظ توضيح ذلك في الشكل (٧ ـ ٢) :

الان بفرض ان $g = x . y$ فان $y = \dfrac{g}{x}$ وان $dy = \dfrac{dg}{x}$. عليه فان

$$F(z) = \int_{-\infty}^{0} \left[\int_{z}^{-\infty} f\left(x, \frac{g}{x}\right) \frac{dg}{x} \right] dx + \int_{0}^{\infty} \left[\int_{-\infty}^{z} f\left(x, \frac{g}{x}\right) \frac{dg}{x} \right] dx$$

$$= A + B$$



الشكل (٧ ـ ٢) : توضيح لايجاد توزيع $Z = X \cdot Y$.

$$= \int_{-\infty}^{z}\left[\int_{-\infty}^{0} f\left(x, \frac{g}{x}\right) \cdot \frac{1}{-x} dx\right] dg + \int_{-\infty}^{z}\left[\int_{0}^{\infty} f\left(x, \frac{g}{x}\right) \cdot \frac{1}{x} \, dx\right] dg$$

$$\therefore F(z) = \int_{-\infty}^{z}\left[\int_{-\infty}^{\infty} \frac{1}{|x|} \cdot f\left(x, \frac{g}{x}\right) dx\right] dg$$

وباشتقاق الطرفين نسبة الى z نحصل على

$$F'(z) = f(z) = \int_{\Omega_x} |x|^{-1} \cdot f\left(x, \frac{z}{x}\right) dx$$

وبنفس الاسلوب يمكن البرهنة ان

$$f(z) = \int_{\Omega_y} |y|^{-1} . f\left(\frac{z}{y}, y\right) dy$$

وفي حالة كون ان x ، y مستقلين فان

$$f(z) = \int_{\Omega_x} |x|^{-1} . f(x) . f\left(\frac{z}{x}\right) dx = \int_{\Omega_y} |y|^{-1} . f\left(\frac{z}{y}\right) . f(y)\, dy$$

اما في حالة قسمة متغيرين فان مسألة استنتاج توزيع $V = \frac{X}{Y}$ لاتختلف كثيراً عن مسألة استنتاج توزيع $Z = X.Y$ سوى انه يتم الفرض بأن $u = \frac{1}{Y}$ وهذا يعني ان $V = X.u$ وكأن ذلك يعني استنتاج التوزيع الاحتمالي لحاصل ضرب المتغيرين X, u.

مثال (١٥) : افرض ان X, Y متغيران عشوائيان مستقلان كل منهما يتوزع وفق دالة التوزيع المنتظم المستمر على الفترة $(0, 1)$. جد التوزيع الاحتمالي لكل من $Z = X.Y$ ، $V = \frac{X}{Y}$.

الحل

$$f(z) = \int_{\Omega_x} |x|^{-1} . f(x) . f\left(\frac{z}{x}\right) dx$$

$$= \int_0^1 \frac{1}{x} . (1) . f\left(\frac{z}{x}\right) dx$$

حيث ان $|x| = x$ في الفترة $(0,1)$. كذلك فان $x = \frac{z}{y}$ وهذا يعني ان $x = z$ عندما $y = 1$ وان $x \to \infty$ عندما $y = 0$. وحيث ان قيم X معرفة على الفترة $(0,1)$ فاذن $z < x < 1$ عليه فان

$$f(z) = \int_z^1 \frac{1}{x} dx = [\ln x]_z^1 = -\ln z, 0 < z < 1$$

كذلك فان

$$f(v) = \int_{\Omega_y} |y| . f(vy, y) dy$$

$$= \int_0^1 |y| . f(vy) . f(y) dy = \int_0^1 y . f(vy) dy$$

وحيث ان $x = vy$ فاذن $y = \frac{x}{v}$ وهذا يعني ان $y = 0$ عندما $x = 0$ وان $y = \frac{1}{v}$ عندما $x = 1$. عليه فان :

$$f(v) = \int_0^1 y\, dy + \int_0^{1/v} y dy$$

$$= \frac{1}{2} \text{ (على الفترة } (0, 1)) + \frac{1}{2v^2} \text{ (على الفترة } (1, \infty))$$

لاحظ من هذا المثال ان :

$$\int_{\Omega_v} f(v) dv = \frac{1}{2} \int_0^1 dv + \frac{1}{2} \int_1^\infty \frac{1}{v^2} dv$$

$$= \frac{1}{2} + \frac{1}{2} \left[-\frac{1}{v} \right]_1^\infty = \frac{1}{2} - \frac{1}{2} (0 - 1) = \frac{1}{2} + \frac{1}{2} = 1$$

٧ ـ ٢ : استنتاج التوزيعات باستخدام الدالة المولدة للعزوم .

في بعض الاحيان نلاقي صعوبة في استنتاج التوزيع الاحتمالي عن طريق الدالة التوزيعية مما يتطلب الامر استخدام اسلوب آخر يمكن من خلاله استنتاج التوزيع . هذا الاسلوب هو استخدام الدالة المولدة للعزوم .

لقد سبق وان ذكرنا في الفقرة (٢ ـ ٢ ـ ١) بانه اذا كانت الدالة المولدة للعزوم موجودة فانها تحدد التوزيع الاحتمالي الذي اشتقت منه والعكس صحيح ايضاً بسبب خاصية وحدانية هذه الدالة . كذلك فقد تم استخدام هذا الاسلوب في فقرات عديدة من الفصلين الخامس والسادس . وفيما يلي وصف لهذا الاسلوب :

افرض ان $X_1, X_2, \ldots, X_n$ متغيرات عشوائية بدالة كثافة احتمالية مشتركة $f(x_1, x_2, \ldots, x_n)$ او كتلة احتمالية مشتركة $P(x_1, x_2, \ldots, x_n)$ ، وافرض ان $Y = g(x_1, x_2, \ldots, x_n)$ دالة بدلالة هذه المتغيرات ، واننا نرغب في استنتاج التوزيع الاحتمالي الى Y . الان بفرض ان الدالة المولدة لعزوم Y موجودة فان ذلك يعني امكانية تعريف هذه الدالة بالشكل :

$$M_Y(t) = Ee^{tY}$$

اي ان

$$M_Y(t) = Ee^{t \cdot g(X_1, X_2, \cdots, X_n)}$$

$$= \int_{x_1} \int_{x_2} \cdots \int_{x_n} e^{t \cdot g(x_1, x_2, \cdots, x_n)} . f(x_1, x_2, \ldots x_n) dx_1 . dx_2 \ldots dx_n$$

في حالة المتغيرات من النوع المستمر

$$= \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} e^{t \cdot g(x_1, x_2, \cdots, x_n)} . P(x_1, x_2, \ldots, x_n)$$

في حالة المتغيرات من النوع المتقطع .

وبعد اجراء عملية التكامل (او الجمع) فاننا سوف نحصل على دالة بدلالة t هذه الدالة تمثل الدالة المولدة لعزوم Y . ومن خلال مقارنة $M_Y(t)$ بدوال توليد العزوم للتوزيعات التي سبق دراستها في الفصلين الخامس والسادس يمكن تحديد التوزيع الاحتمالي المقابل الى $M_Y(t)$ طالما ان هذه الدالة تتصف بصفة الوحدانية وانها تحدد التوزيع الاحتمالي الذي يقابلها بشكل متكامل . علماً اننا في بعض الاحيان قد نحصل على دالة مولدة للعزوم الا انها لاتشبه تلك الدوال التي

سبق لنا دراستها وذلك لايعني ان التوزيع الاحتمالي المطلوب غير ممكن التحديد بل ان ذلك ممكن طالما ان الدالة المولدة للعزوم موجودة . ويمكن التوصل لذلك باستخدام الدالة المميزة التي سبق التنويه عنها في الفقرة (٢ ـ ٣) .

مثال (١٦) : افرض ان X متغير عشوائي ذو توزيع طبيعي معياري . جد التوزيع الاحتمالي الى $Y = X^2$

الحل : نفرض ان الدالة المولدة لعزوم Y موجودة . فاذن

$$M_Y(t) = Ee^{tY} = Ee^{tX^2}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx^2} . e^{-\frac{1}{2}x^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2(1-2t)} dx$$

$$= (1-2t)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\ (1-2t)^{-\frac{1}{2}}} . e^{-\frac{1}{2}x^2(1-2t)} dx$$

لكن قيمة التكامل في الصيغة الاخيرة مساوية للواحد نظراً لان التكامل يجري على دالة توزيع طبيعي وسطه صفر وتباينه $(1-2t)^{-1}$, $t < \frac{1}{2}$. فاذن ،

$$M_Y(t) = (1-2t)^{-\frac{1}{2}}, t < \frac{1}{2}$$

والدالة الاخيرة تمثل الدالة المولدة لعزوم توزيع كاما بالمعلمتين $\alpha = \frac{1}{2}$, $\beta = 2$ وهذا يعني ان $Y = X^2 \sim G\left(\frac{1}{2}, 2\right)$.

مثال (١٧) : افرض ان X متغير عشوائي يتوزع وفق دالة توزيع ثنائي الحدين بالمعلمتين n, P . جد التوزيع الاحتمالي الى $Y = n - X$.

الحل : نفرض ان الدالة المولدة لعزوم Y موجودة . فاذن

$$M_Y(t) = Ee^{tY} = Ee^{t(n-X)}$$
$$= e^{nt}.M_X(-t)$$

وحيث ان X هو ذا توزيع ثنائي الحدين . فاذن

$$M_X(-t) = (q + Pe^{-t})^n, q = 1 - P$$

عليه فان

$$M_Y(t) = (e^t)^n.(q + Pe^{-t})^n$$
$$\therefore M_Y(t) = (qe^t + P)^n$$

والصيغة الآخيرة تمثل الدالة المولدة لعزوم توزيع ثنائي الحدين بالمعلمتين n, q . فاذن $Y = n - X \sim b(n, q)$.

٧ ــ ٤ : استنتاج التوزيعات باستخدام التحويلات .

نلاحظ في بعض الاحيان صعوبة في استنتاج التوزيعات الاحتمالية باستخدام مفهوم الدالة التوزيعية او الدالة المولدة للعزوم مما يستوجب الامر التوجه نحو اسلوب آخر لهذه الحالة او تلك . هذا الاسلوب هو استخدام التحويلات transformations في استنتاج التوزيع الاحتمالي . لقد سبق وان استخدمنا هذا الاسلوب في الفقرة (٦ ــ ٢) لدى دراستنا لموضوع التوزيع الطبيعي . وفيما يلي وصف لهذا الاسلوب :

٧ ــ ٤ ــ ١ : استنتاج التوزيعات المتقطعة باستخدام التحويلات .

افرض ان X متغير عشوائي متقطع بدالة كتلة احتمالية P(x) معرف على فضاء متقطع مثل Ω_x . وافرض ان $y = g(x)$ يقال في هذه الحالة ان $y=g(x)$ يمثل تحويل من X الى Y وان هذا التحويل يُطبق maps فضاء Ω_x , X فوق فضاء Y . ليكن Ω_y . وان Ω_y ناتج في الحقيقة من خلال تحويل اية قيمة

معرفة في Ω_x استناداً الى $y = g(x)$ كما ويستوجب هذا التحويل ان يقابل كل قيمة معرفة في Ω_x قيمة واحدة فقط في Ω_y كذلك لكل قيمة معرفة في Ω_y يقابلها قيمة واحدة فقط في Ω_x . وعندئذٍ يقال ان هنالك تقابلاً (واحد لواحد one – to – one) لعنصر في Ω_x مع عنصر في Ω_y استناداً للتحويل $y = g(x)$ وفق هذا التصور يمكن الاستنتاج بان Y دالة وحيدة القيمة بدلالة X وان X دالة وحيدة القيمة بدلالة Y . وهذا يعني ان X دالة معكوسة الى Y ولتكن هذه الدالة $x = W(y)$، عندئذٍ اذا كانت $y \in \Omega_y$ فان $x = W(y) \in \Omega_x$ ، وهذا يعني ان الحادثتان $\{Y = y\}$, $\{X = W(y)\}$ متكافئتين الامر الذي يمكن من خلاله استنتاج التوزيع الاحتمالي الى Y . فاذن

$$P(y) = P_r(Y = y) = P_r(X = W(y)) = P(W(y)), y \in \Omega_y$$

اما في حالة دوال الكتلة الاحتمالية المشتركة فان الامر لايختلف كثيراً من حيث المضمون . فعلى فرض ان $P(x_1, x_2)$ تمثل دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 معرفتين على فضاء ثنائي مثل $\Omega_{x_1x_2}$ ، وافرض ان $y_1 = g_1(x_1, x_2)$, $y_2 = g_2(x_1, x_2)$ يمثل كل منهما تحويل يُطبق $\Omega_{x_1x_2}$ فوق $\Omega_{y_1y_2}$ بحيث ان لكل زوج مثل (x_1, x_2) هنالك زوج واحد فقط مثل (y_1, y_2) معرف في $\Omega_{y_1y_2}$ وان الدالة المعكوسة الى كل من y_1, y_2 هي $x_1 = W_1(y_1, y_2)$, $x_2 = W_2(y_1, y_2)$. واستناداً لما تقدم فان دالة الكتلة الاحتمالية المشتركة الى Y_1, Y_2 هي :

$$\begin{aligned} P(y_1, y_2) &= P_r(Y_1 = y_1, Y_2 = y_2) = P_r(X_1 = W_1(y_1, y_2), X_2 = W_2(y_1, y_2)) \\ &= P(W_1(y_1, y_2), W_2 = (y_1, y_2)) \end{aligned}$$

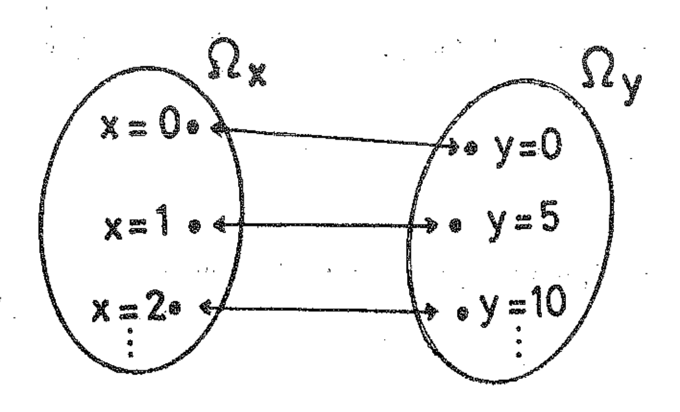
ويمكن تعميم ماتقدم لحالة وجود اكثر من متغيرين .

وتجدر الاشارة هنا انه في بعض الاحيان يصادفنا تحويل واحد فقط بدلالة متغيرين او اكثر . فمثلاً للحالة السابقة وبفرض ان التحويل المعطى هو $y_1 = g_1(x_1, x_2)$ وتطلب الامر استنتاج توزيع Y_1 فان ذلك يستلزم منا تعريف تحويل آخر بدلالة احد المتغيرين او كلا المتغيرين X_1, X_2 ان كان ذلك ممكناً وليكن $y_2 = g_2(x_1, x_2)$. عندئذٍ فان التحويل الجديد y_1 يكمل التحويل y_2 لفرض تحقيق التقابل بين ازواج القيم المعرفة في $\Omega_{x_1x_2}$ مع تلك المعرفة في

$\Omega_{y_1 y_2}$ وبذلك يمكن استنتاج دالة الكتلة الاحتمالية المشتركة الى Y_1, Y_2 ومن ثم يصار الى ايجاد الدالة الحدية للمتغير Y_1 من خلال اجراء عملية الجمع لدالة الكتلة الاحتمالية المشتركة حول فضاء Y_2 . نستشف مما تقدم ان عدد التحويلات المطلوبة لفرض استنتاج دالة الكتلة الاحتمالية المشتركة يجب ان يكون بنفس عدد المتغيرات التي يتضمنها التوزيع المشترك . وفيما يلي بعض الامثلة التي توضح ماسبق .

مثال (١٨) : افرض ان X متغير عشوائي ذو توزيع پواسون بالمعلمة m . جد التوزيع الاحتمالي الى $Y = 5X$.

الحل : واضح ان $\Omega_x = \{ x : x = 0, 1, 2, \ldots \}$ وان $P(x) > 0$ لقيم x المعرفة في Ω_x . ان التحويل $y = 5x$ يُطبق فضاء X على فضاء Y المعرف بالمجموعة $\Omega_y = \{ y : y = 0, 5, 10, \ldots \}$ الناتجة من تحويل كل قيمة معرفة في Ω_x استناداً الى التحويل $y = 5x$ وان لكل قيمة معرفة في Ω_x يقابلها قيمة واحدة فقط في Ω_y ، وهذا يعني ان هنالك تقابلاً (واحد لواحد) لعنصر في Ω_x مع عنصر في Ω_y . لاحظ الشكل (٧ ــ ٣) .



الشكل (٧ ــ ٣) ، توضيح لتقابل عناصر Ω_x مع عناصر Ω_y .

كذلك فان الدالة المعكوسة هي$x = \frac{1}{5}y$ عليه فان

$$P(y) = P_r(Y = y) = P_r\left(X = \frac{1}{5}y\right) = \frac{m^{y/5}.e^{-m}}{(y/5)!} \; ;y = 0, 5, 10, \ldots$$

مثال (١٩) : افرض ان $X \sim b\left(4, \frac{3}{4}\right)$ جد دالة الكتلة الاحتمالية الى$Y = X^2$

الحل : واضح ان $\Omega_x = \{x: x = 0, 1, 2, 3, 4\}$ وان $P(x) > 0$ لقيم x المعرفة في Ω_x .

ان التحويل $y = x^2$ يطبق فضاء X على فضاء Y المعرف بالمجموعة $\Omega_y = \{y: y = 0, 1, 4, 9, 16\}$ الناتجة من تحويل كل قيمة معرفة في Ω_x استناداً للتحويل $y = x^2$ الا انه وبشكل عام التحويل$y = x^2$ لايحدد تقابل عنصر في Ω_x مع عنصر في Ω_y بسبب ان الدالة المعكوسة هي $x = \pm\sqrt{y}$ الا انه وبسبب كون فضاء X لايتضمن قيم سالبة لهذا المتغير لذا فان الاشارة السالبة تهمل وتكون الدالة المعكوسة هي $x = \sqrt{y}$. وعندئذ :

$$P(y) = P_r(Y = y) = P_r(X = \sqrt{y}) = C^4_{\sqrt{y}} \left(\frac{3}{4}\right)^{\sqrt{y}} \cdot \left(\frac{1}{4}\right)^{4-\sqrt{y}}$$

$; y = 0, 1, 4, 9, 16$

مثال (٢٠) : افرض ان $X_1 \sim Po(m_1)$ مستقل عن $X_2 \sim Po(m_2)$ جد التوزيع الاحتمالي الى $Y = X_1 + X_2$.

الحل : واضح ان التحويل Y يتضمن المتغيرين X_1, X_2 مما يتطلب الامر تعريف تحويل آخر مكمل كي يكون عدد التحويلات مساوٍ لعدد المتغيرات . ولنفرض ان $Z = X_2$. وهذا يعني ان $Z \sim Po(m_2)$. ان دالة الكتلة الاحتمالية المشتركة للمتغيرين X_1, X_2 هي

$$P(x_1, x_2) = \frac{m_1^{x_1} . m_2^{x_2} . e^{-(m_1 + m_2)}}{x_1! . x_2!} \quad ; x_1, x_2 = 0, 1, 2, \dots$$

الان

$$y = x_1 + x_2 \quad , \quad z = x_2$$

فاذن الدوال المعكوسة هي :

$$x_2 = z \quad , \quad x_1 = y - z \quad , \quad z < y$$

وهذا يعني ان

$$y = 0, 1, 2, \dots \quad , z = 0, 1, 2, \dots, y$$

فاذن

$$P(y, Z) = P_r(Y = y, Z = z) = P_r(x_1 = y - Z, x_2 = z)$$

$$= \frac{m_1^{y-z} . m_2^{z} . e^{-(m_1 + m_2)}}{(y - z)! . z!}$$

وعندئذٍ فان

$$P(y) = \sum_{z=0}^{y} P(y, z) = e^{-(m_1 + m_2)} \sum_{z=0}^{y} \frac{m_1^{y-z} . m_2^{z}}{(y - z)! . z!}$$

$$= \frac{e^{-(m_1 + m_2)}}{y!} \sum_{z=0}^{y} \frac{y!}{z!(y - z)!} . m_1^{y-z} . m_2^{z}$$

$$= \frac{e^{-(m_1 + m_2)}}{y!} \sum_{z=0}^{y} C_z^y m_2^z . m_1^{y-z}$$

$$\therefore P(y) = \frac{e^{-(m_1 + m_2)} . (m_1 + m_2)^y}{y!} \quad ; y = 0, 1, 2, \dots$$

لاحظ ان الدالة الاخيرة تمثل دالة توزيع پواسون بالمعلمة $(m_1 + m_2)$ فاذن $Y = X_1 + X_2 \sim Po(m_1 + m_2)$.

٧ ـ ٤ ـ ٢ : ١ استنتاج التوزيعات المستمرة باستخدام التحويلات .

افرض ان X متغير عشوائي مستمر بدالة كثافة احتمالية $f(x)$ معرف على فضاء مستمر مثل Ω_x بحيث ان $f(x) > 0$ لجميع قيم $x \in \Omega_x$ وافرض ان $y = g(x)$ يمثل تحويلاً من X الى Y بحيث ان هذا التحويل يُطبق Ω_x فوق Ω_y الناتج من تحويل اية قيمة معرفة في Ω_x استناداً الى $y = g(x)$ وان هنالك تقابلاً (واحد لواحد) مابين Ω_y , Ω_x بحيث ان كل عنصر في Ω_x يقابله عنصر واحد في Ω_y وان كل عنصر في Ω_y يقابله عنصر واحد فقط في Ω_x . وافرض ان الدالة المعكوسة الى $y = g(x)$ هي $x = w(y)$ وان المشتقة $dx / dy = w'(y)$ موجودة ومستمرة عند جميع قيم $y \in \Omega_y$ ، وطبقاً لما تقدم يمكن استنتاج دالة التوزيع الاحتمالي الى Y وفقاً للاتي

$f(y) = f(x)]_{x=w(y)} \cdot |w'(y)| \; ; y \in \Omega_y$ حيث $|w'(y)|$ تعني القيمة المطلقة لمشتقة الدالة $w(y)$ والتي غالباً مايطلق عليها بمعامل معكوس التحويل $x = w(y)$ اومعامل التحويل او جاكوبيان Jacobian التحويل ويرمز لذلك عادة بالرمز J اي ان $J = w'(y)$. وعندئذ فان $f(y) = f(x)]_{x=w(y)} \cdot |J|$.

مثال (٢١) : افرض ان X متغير عشوائي يتوزع وفق دالة التوزيع المنتظم المستمر على الفترة $(0, 1)$. جد دالة الكثافة الاحتمالية الى $Y = -2\ln X$.

الحل : واضح من هذا المثال ان $y = g(x) = -2\ln x$ فاذن الدالة المعكوسة هي $x = w(y) = e^{-\frac{1}{2}y}$ وان هنالك تقابل بين عناصر Ω_x المعرفة بالفترة $(0,1)$ وعناصر Ω_y المعرفة بالفترة $(0, \infty)$ الناتجة طبقاً للتحويل $y = -2\ln x$ بحيث ان كل عنصر في Ω_x يقابله عنصر واحد فقط في Ω_y وكل عنصر في Ω_y يقابله عنصر واحد فقط في Ω_x لاحظ الشكل (٧ ـ ٤) وان .

$$J = \frac{dx}{dy} = w'(y) = -\frac{1}{2} e^{-\frac{1}{2}y}$$

الشكل (٧ ـ ٤) : توضيح لتقابل عناصر Ω_x وعناصر Ω_y

لاحظ ان المشتقة موجودة ومستمرة لكافة قيم $y \in \Omega_y$. فاذن

$$|J| = \frac{1}{2} e^{-\frac{1}{2} y}$$

عليه فان

$$f(y) = f(x)\Big]_{x = e^{-\frac{1}{2} y}} \cdot |J|$$

$$= (1) . \frac{1}{2} e^{-\frac{1}{2} y} = \frac{1}{2} e^{-\frac{1}{2} y} , y \geq 0$$

والدالة الاخيرة تمثل دالة التوزيع الاسي بالمعلمة $\theta = \frac{1}{2}$ فاذن

$$Y = -2\ln X \sim EXP\left(\frac{1}{2}\right)$$

مثال (٢٢): افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = 2xe^{-x^2} ; x > 0$ جد التوزيع الاحتمالي الى $Y = X^2$

الحل : ان $y = x^2$, $0 < y < \infty$ وان $x = \pm \sqrt{y}$. وحيث ان قيم x محددة بالفترة $(0, \infty)$ فاذن $x = \sqrt{y}$ وان $J = \frac{1}{2} y^{-\frac{1}{2}}$

فاذن

$$f(y) = f(x)\Big]_{x=\sqrt{y}} \cdot |J| = 2y^{\frac{1}{2}} \cdot e^{-y} \cdot \frac{1}{2} y^{-\frac{1}{2}}$$

$$= e^{-y} \ , \ y \geq 0$$

ويلاحظ هنا ان

$$.Y = X^2 \sim EXP(1)$$

ووفق نفس المفهوم اعلاه يمكن استنتاج التوزيع الاحتمالي المشترك لحالة وجود اكثر من متغير واحد. فعلى فرض ان $f(x_1, x_2)$ تمثل دالة الكثافة الاحتمالية المشتركة للمتغيرين X_1, X_2 وان كل من $y_1 = g_1(x_1, x_2)$ و $y_2 = g_2(x_1, x_2)$ يمثل تحويل من X_1, X_2 الى Y_1, Y_2 بحيث ان هذين التحويلين يعرفان تقابل زوج واحد فقط من $\Omega_{x_1 x_2}$ مع زوج واحد فقط في $\Omega_{y_1 y_2}$ اي مانعنيه ان التحويلان y_1, y_2 يطبقان الفضاء الثنائي $\Omega_{x_1 x_2}$ على الفضاء الثنائي $\Omega_{y_1 y_2}$. وافرض ان الدالة المعكوسة الى x_1 هي $x_1 = w_1(y_1, y_2)$ والى x_2 هي $x_2 = w_2(y_1, y_2)$. ان معامل التحويل في هذه الحالة يتمثل بمحدد مصفوفة ذات مرتبة 2×2 عناصر الصف الاول منها تمثل المشتقة الجزئية الى x_1 نسبة الى y_1 و y_2 في حين ان عناصر الصف الثاني منها تمثل المشتقة الجزئية الى x_2 نسبة الى y_1 و y_2 اي ان :

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$$

$$= \frac{\partial x_1}{\partial y_1} \cdot \frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2} \cdot \frac{\partial x_2}{\partial y_1}$$

وبذلك فان

$$f(y_1, y_2) = f(x_1, x_2)\Big]_{\substack{x_1 = w_1(y_1, y_2) \\ x_2 = w_2(y_1, y_2)}} \cdot |J|$$

وفي حالة كون ان $X_1, X_2, \dots, X_k$ متغيرات عشوائية مستمرة بدالة كثافة احتمالية مشتركة $f(x_1, x_2, \dots, x_k)$ وان $i = 1, 2, \dots$ و $y_i = g_i(x_1, x_2, \dots, x_k)$ تمثل تحويلات بدلالة هذه المتغيرات تؤدي الى تقابل واحد لواحد بين $\Omega_{x_1 x_2 \dots x_k}$ و $\Omega_{y_1 y_2 \dots y_k}$

وان الدوال المعكوسة هي $x_i = w_i(y_1, y_2, \dots, y_k)$ عندئذ فان معامل التحويل سوف يتمثل بمحدد مصفوفة ذات مرتبة $k \times k$ عناصرها تمثل مشتقات جزئية للمتغيرات القديمة (X_i) نسبة الى المتغيرات الجديدة (Y_i) اي ان :

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_k} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_k}{\partial y_1} & \frac{\partial x_k}{\partial y_2} & \cdots & \frac{\partial x_k}{\partial y_k} \end{vmatrix}$$

وبذلك فان

$$f(y_1, y_2, \dots, y_k) = f(x_1, x_2, \dots, x_k)] \cdot |J|$$

$$\begin{aligned} x_1 &= w_1(\dots) \\ x_2 &= w_2(\dots) \\ &\vdots \\ x_k &= w_k(\dots) \end{aligned}$$

كذلك يجب ان يكون عدد التحويلات مساو لعدد المتغيرات الاصلية اي k. واذا كان عددها اقل من k فان الامر يستوجب تعريف تحويلات اخرى مكملة وبنفس الاسلوب الذي تم التنويه عنه في الفقرة (٧ ـ ٤ ـ ١) .

مثال (٢٣) : افرض ان X_1 , X_2 متغيران عشوائيان مستقلان كل منهما يتوزع وفق دالة التوزيع المنتظم المستمر على الفترة $(0,1)$ جد دالة الكثافة الاحتمالية المشتركة الى Y_1 , Y_2 بحيث ان $Y_1 = X_1 + X_2$ وان $Y_2 = X_1 - X_2$

الحل : نجد اولاً الدالة المعكوسة الى X_1 و X_2 وكالاتي :

$$\left.\begin{aligned} Y_1 &= X_1 + X_2 \\ Y_2 &= X_1 - X_2 \end{aligned}\right\} \rightarrow X_1 = \frac{1}{2}(Y_1 + Y_2)$$

وان

$$\left.\begin{aligned} Y_1 &= X_1 + X_2 \\ Y_2 &= X_1 - X_2 \end{aligned}\right\} \rightarrow X_2 = \frac{1}{2}(Y_1 - Y_2)$$

عليه فان

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{-1}{2} \end{vmatrix} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}$$

$$\therefore |J| = \frac{1}{2}$$

وان $f(x_1, x_2) = f(x_1) . f(x_2) = 1$

فاذن

$$f(y_1, y_2) = f(x_1, x_2)] \quad . |J| = (1) . \frac{1}{2} = \frac{1}{2}$$

$$x_1 = \frac{1}{2}(y_1 + y_2)$$

$$x_2 = \frac{1}{2}(y_1 - y_2)$$

تبقى امامنا مشكلة واحدة وهي تحديد فضاء y_1 وفضاء y_2 . لاحظنا ان الدالة المعكوسة الى X_1 كانت $x_1 = \frac{1}{2}(y_1 + y_2)$ وتلك الى X_2 كانت $x_2 = \frac{1}{2}(y_1 - y_2)$.

فاذن عندما :

$$x_1 = 0 \rightarrow \frac{1}{2}(y_1 + y_2) = 0 \rightarrow y_1 = -y_2 \rightarrow y_2 = -y_1$$

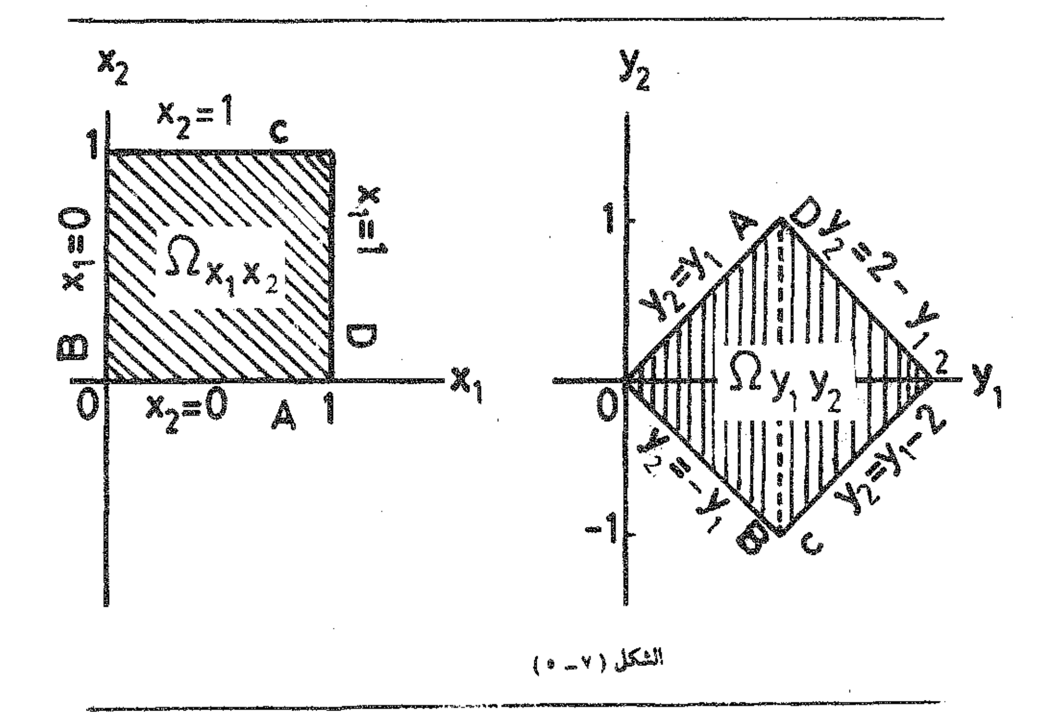
$$x_1 = 1 \rightarrow \frac{1}{2}(y_1 + y_2) = 1 \rightarrow y_1 = 2 - y_2 \rightarrow y_2 = 2 - y_1$$

$$x_2 = 0 \rightarrow \frac{1}{2}(y_1 - y_2) = 0 \rightarrow y_1 = y_2 \rightarrow y_2 = y_1$$

$$x_2 = 1 \rightarrow \frac{1}{2}(y_1 - y_2) = 1 \rightarrow y_1 = 2 + y_2 \rightarrow y_2 = y_1 - 2$$

كذلك فان التحويل $y_1 = x_1 + x_2$ يؤدي الى ان $0 < y_1 < 2$ طالما ان $0 < x_1, x_2 < 1$ وهذا يعني ان $-1 < y_2 < 1$ وكما موضح في الشكل (٧ - ٥) :

لاحظ انه لقيم y_1 في الفترة $(0, 1)$ فان قيم y_2 ستكون معرفة في الفترة $(-y_1, y_1)$ ولقيم y_1 في الفترة $(1, 2)$ فان قيم y_2 ستكون معرفة في الفقرة $(y_1 - 2, 2 - y_1)$.



الشكل (٧ ـ ٥)

مثال (٢٤) افرض ان $X_1 \sim N(0,1)$ مستقل عن $X_2 \sim N(0,1)$. جد التوزيع الاحتمالي الى $Y_1 = X_1/X_2$.

الحل : ان

$$f(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}; -\infty < x_1, x_2 < \infty$$

وان $y_1 = \frac{x_1}{x_2}$. وافرض ان $y_2 = x_2$ فاذن $-\infty < y_2 < \infty$.
ان الدالة المعكوسة الى X_2 هي $x_2 = y_2$ والدالة المعكوسة الى X_1 هي $x_1 = y_1 \cdot y_2$ عليه فان معامل التحويل J هو :

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2 \quad \therefore |J| = y_2$$

$$\therefore \quad f(y_1, y_2) = f(x_1, x_2)\Big]_{\substack{x_1 = y_1 y_2 \\ x_2 = y_2}} \cdot |J|$$

$$= \frac{1}{2\pi} e^{-\frac{1}{2}(y_1^2 y_2^2 + y_2^2)} \cdot y_2 = \frac{y_2}{2\pi} e^{-\frac{1}{2} y_2^2 (y_2^2 + 1)} \quad , -\infty < y_1, y_2 < \infty$$

$$\therefore f(y_1) = \int_{-\infty}^{\infty} \frac{y_2}{2\pi} e^{-\frac{1}{2} y_2^2 (y_1^2 + 1)} dy_2$$

$$= \int_0^{\infty} \frac{y_2}{\pi} e^{-\frac{1}{2}(y_2^2 (y_1^2 + 1)} dy_2$$

$$= -\frac{1}{\pi(y_1^2 + 1)} \cdot \left[e^{-\frac{1}{2} y_2^2 (y_1^2 + 1)} \right]_0^{\infty}$$

$$= -\frac{1}{\pi(y_1^2 + 1)} [0 - 1] = \frac{1}{\pi(1 + y_1^2)} ; -\infty < y_1 < \infty$$

والدالة الاخيرة تمثل دالة توزيع كوشي بالمعلمتين $a = 0$, $b = 1$.

تمارين الفصل السابع

٧ ــ ١ : اذا علمت ان $X_1 \sim N(2,4)$, $X_2 \sim N(3,9)$, $X_3 \sim N(4,16)$
وان $\rho_{12} = 0.6$, $\rho_{13} = 0.4$, $\rho_{23} = -0.7$. جد
مايلي :
أ ــ الوسط والتباين الى $Y = X_1 + X_2 + X_3$, $Y = 2X_1 - 3X_2 - 4X_3$.
ب ــ الوسط والتباين الى $Y = X_1 . X_2$, $Z = X_1 . X_3$, $V = X_2 . X_3$.
جـ ــ الوسط والتباين الى $Y = X_1 / X_2$, $Z = X_1 / X_3$, $V = X_2 / X_3$.

٧ ــ ٢ : اذا علمت ان $X_1 \sim \text{beta}(2,4)$ مستقل عن $X_2 \sim \text{beta}(4,6)$ جد
الوسط والتباين الى $Y = X_1 / X_2$

٧ ــ ٣ : افرض ان θ متغير عشوائي يتوزع وفق دالة التوزيع المنتظم المستمر على
الفترة $(0, 2\pi)$ وافرض ان $Y = A \cos\theta$ حيث ان A ثابت
حقيقي .
جد التوزيع الاحتمالي الى Y باستخدام اسلوب الدالة التوزيعية .

٧ ــ ٤ : افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = e^{-x}, x \geq 0$
جد التوزيع الاحتمالي الى $Y = 1 - e^{-x}$ باستخدام اسلوب الدالة التوزيعية .

٧ ــ ٥ : افرض ان X,Y متغيران عشوائيان مستقلان كل منهما يتوزع وفق دالة
التوزيع الاسي بالمعلمة $\theta = 1$. جد دالة الكثافة الاحتمالية الى
$Z = \sqrt{X + Y}$ (ملاحظة : افرض ان $x = v^2$)

٧ ــ ٦ : افرض ان X,Y متغيران عشوائيان مستقلان بحيث ان $f(x) = xe^{-\frac{1}{2}x^2}$; $x > 0$
وان $f(y) = ye^{-\frac{1}{2}y^2}$; $y > 0$. بين ان دالة الكثافة الاحتمالية
للمتغير $Z = X / Y$ هي

$$f(Z) = \frac{2Z}{(Z^2 + 1)}, Z > 0$$

٧ ـ ٧ : اذا علمت ان X متغير عشوائي بدالة كثافة احتمالية $f(x) = \frac{6x}{(1+x)^4}$

$x > 0$ ؛ بين ان دالة الكثافة الاحتمالية الى $Y = \frac{1}{X}$ هي $(y) = \frac{6y}{(1+y)^4}, y > 0$

٧ ـ ٨ : استخدم اسلوب الدالة التوزيعية في استنتاج التوزيع الاحتمالي الى $Y = e^{-\theta X}, \theta > 0$

علماً ان $f(x) = \frac{1}{x^2}, x \geq 1$

٧ ـ ٩ : افرض ان $X \sim N(\mu, \sigma^2)$. جد التوزيع الاحتمالي الى $Y = e^X$

٧ ـ ١٠ : اذا علمت ان الدالة التوزيعية لمتغير عشوائي X $- \infty < x < \infty$

هي $F(x) = \text{Exp}\left(-\text{EXP}\left[-\left(\frac{x-\alpha}{\beta}\right)\right]\right)$ ، α, β معالم

بحيث ان $- \infty < \alpha < \infty, \beta > 0$. بين ان التوزيع الاحتمالي الى

$Y = e^{-\frac{x-\alpha}{\beta}}$ هو توزيع اسي بالمعلمة $\theta = 1$.

٧ ـ ١١ : اذا علمت ان $X \sim \text{beta}(\alpha, \beta)$ ، جد التوزيع الاحتمالي الى $Y = 1 - X$

٧ ـ ١٢ : استخدم اسلوب الدالة المولدة للعزوم في استنتاج التوزيع الاحتمالي الى

$Y = X_1 + X_2 + \ldots + X_n$ لكل حالة من الحالات التالية علماً ان

المتغيرات $X_1, X_2, , X_n$. مستقلة تصادفياً :

أ ـ ان $X_i \sim G(\alpha_i, \beta)$

ب ـ ان X_i يتوزع كتوزيع هندسي بالمعلمة P

جـ ـ ان X_i يتوزع كتوزيع ثنائي الحدين السالب بالمعلمتين r_i, P

د ـ ان X_i يتوزع كتوزيع ثنائي الحدين بالمعلمتين n_i, P

٧ ـ ١٣ : اذا علمت ان X يتوزع وفق دالة توزيع منتظم مستمر على الفترة $(0,1)$

جد التوزيع الاحتمالي الى $Y = X^{-1}$

٧ ـ ١٤ : اذا علمت ان $X_1 \sim b(n_1, P)$ مستقل عن $X_2 \sim b(n_2, p)$ استخدم

الاسلوب الموضح في الفقرة (٧ ـ ٤ ـ ١) لايجاد توزيع $Y = X_1 + X_2$

٧ ـ ١٥ افرض ان $X_1 \sim N(\mu_1, \sigma_1^2)$ مستقل عن $X_2 \sim N(\mu_2, \sigma_2^2)$
استخدم الاسلوب الموضح في الفقرة (٧ ـ ٢ ـ ١) لايجاد توزيع $Y = X_1 - X_2$

٧ ـ ١٦ افرض ان $X \sim EXP(\theta)$ مستقل عن $Y \sim EXP(\theta)$ برهن ان
$g = \dfrac{X}{X+Y}$ يتوزع كتوزيع منتظم مستمر على الفترة $(0, 1)$.

٧ ـ ١٧ اذا كانت $Z_1, Z_2, \ldots, Z_n$ متغيرات عشوائية مستقلة كل منها
يتوزع كتوزيع $N(0,1)$ برهن ان

$$\frac{1}{2} \sum_{i=1}^{n} Z_i^2 \sim G\left(\frac{n}{2}, 1\right)$$

٧ ـ ١٨ اذا كانت $X_1, X_2, \ldots X_n$ متغيرات عشوائية مستقلة بحيث ان
$X_i \sim G\left(\frac{1}{n}, 1\right)$ برهن ان $\sum_{i=1}^{n} X_i \sim EXP(1)$.

٧ ـ ١٩ اذا كان $X \sim F(\alpha, \beta)$ جد التوزيع الاحتمالي للمتغير $Y = \dfrac{X}{1-X}$

٧ ـ ٢٠ افرض ان $X_1 \sim G(\alpha, 1)$ مستقل عن $X_2 \sim G(\beta, 1)$. برهن ان
$Y = X_1/(X_1 + X_2)$ يتوزع كتوزيع $B(\alpha, \beta)$.





المعاينة والتوزيعات المقيدة

الفصل الثامن
المعاينة والتوزيعات المقيدة
Sampling & limiting distributions

استعرضنا في الفصلين الخامس والسادس اهم التوزيعات الاحتمالية الشائعة الاستخدام في النظرية الاحصائية من خلال دراستنا لأهم خصائص هذه التوزيعات وعلاقة بعضها بالبعض الاخر. في حين اختص الفصل السابع بدراسة توزيعات دوال المتغيرات العشوائية من حيث توقعاتها وخصائصها فضلاً عن استعراض اهم الطرق التي من خلالها يمكن استنتاج التوزيع الاحتمالي لهذه الدوال.

في هذا الفصل سوف نركز الاهتمام على دراسة مفهوم المعاينة ومفهوم التوزيعات المقيدة والتقارب التصادفي ودورهما في استنتاج التوزيعات الاحتمالية على الرغم من استخدامنا لهذين المفهومين بشكل او بآخر غير مباشر في بعض فقرات الفصلين الخامس والسادس.

٨ ـ ١ : المعاينة Sampling

يقصد بالمعاينة « اسلوب » او « طريقة » يمكن بواسطتها الحصول على « عينة Sample » من المفردات units من مجتمع population معين. ويقصد بالمجتمع (او المجتمع الاحصائي) بانه كافة المفردات التي تشترك بخاصية (او مجموعة خصائص) معينة. على سبيل المثال مجتمع طلبة وطالبات جامعة الموصل، مجتمع الاسر الساكنة في مركز مدينة الموصل، وغيرها من الامثلة. وهذا يعني ان المجتمع الاحصائي يمثل جمع من المفردات ذات خاصية (او مجموعة خصائص) معينة مشتركة تخص دراسة معينة. في حين يقصد بالعينة بانها مجموعة من المفردات تشكل جزء (مجموعة جزئية) من المجتمع الاحصائي يتم اختيارها وفق قواعد واصول معينة تسمى اساليب المعاينة Sampling techniques ومن اهم هذه الاساليب

أ ــ المعاينة العشوائية البسيطة Simple random sampling

ب ــ المعاينة العشوائية الطبقية Stratified random sampling

جـ ــ المعاينة العشوائية المنتظمة Systematic random sampling

ونظراً لكون الطالب قد سبق وان درس هذه الاساليب وغيرها بشكل مفصل فاننا سوف لن ندخل في تفاصيلها كي لانخرج عن نطاق هذا الكتاب .

٨ ــ ١ ــ ١ : المعاينة العشوائية Random Sampling

يقصد بالمعاينة العشوائية عملية سحب عينة من المفردات من مجتمع احصائي بالشكل الذي يضمن لكل مفردة من مفردات المجتمع نفس الفرصة في الاختيار لان تكون واحدة من مفردات تلك العينة . فعلى فرض ان المجتمع الاحصائي محدود وعدد مفرداته هو N وتطلب الامر سحب عينة قوامها n مفردة ، $n < N$ ، من هذا المجتمع فان الاختيار العشوائي لمفردات هذه العينة يضمن احتمالاً قدره $\frac{1}{N}$ كفرصة لاختيار اية مفردة من مفردات المجتمع دون ان يكون هنالك اي مبرر (تحيز) لاختيار هذه المفردة دون الاخرى . كذلك فان عدد العينات الممكنة الاختيار من هذا المجتمع هو C_n^N . ان اسلوب المعاينة الذي يضمن نفس الفرصة في اختيار اية مفردة دون اي تحيز يسمى اسلوب معاينة عشوائية في حين ان العينة التي يستحصل عليها وفق هذا الاسلوب تسمى عينة عشوائية random sample .

وعلى فرض ان $X_1 , X_2 , \ldots X_n$ تمثل متغيرات عشوائية بدالة كثافة احتمالية مشتركة $f(x_1 , x_2 , \ldots , x_n)$ (او كتلة احتمالية مشتركة) ، فاذا امكن لنا صياغة التوزيع المشترك بالشكل $f(x_1 , x_2 , \ldots , x_n) = \prod_{i=1}^{n} f(x_i)$

بحيث ان $f(x_i) = f(x) \quad , \forall i = 1, 2, \ldots, n$ فان

ذلك يعني ان هذه المتغيرات مستقلة تصادفياً وانها تتوزع وفق نفس التوزيع الاحتمالي المعرف بالدالة $f(x)$. وعندئذٍ يمكن التعبير عن $X_1 , X_2 , \ldots , X_n$

على انها قياسات مفردات عينة عشوائية قوامها n مسحوبة من مجتمع ذو دالة كثافة احتمالية $f(x)$. وهذا يعني ان القياسات $X_1, X_2 , \ldots , X_n$ يمكن النظر اليها على انها متغيرات عشوائية مستقلة كل منها بدالة كثافة احتمالية $f(x)$. اي ان الاصطلاح « عينه عشوائية » مكافيء للاصطلاح « متغيرات عشوائية مستقلة » .

٨ ـ ١ ـ ٢ : المؤشر الاحصائي والمعلمة Parameter and statistic

لاحظنا لدى دراستنا للتوزيعات الاحتمالية في الفصلين الخامس والسادس ان اي توزيع منها عبارة عن عائلة توزيعات كل عضو منها يتحدد من خلال تخصيص قيمة عددية لمعلمة (او معالم) ذلك التوزيع ، واعتبرنا هذه المعلمة (المعالم) كمية (او كميات) ثابتة . الا انه ومن الناحية العملية غالباً ماتكون هذه المعالم مجهولة القيمة . فمثلاً لاحظنا في التوزيع الطبيعي ان كل من μ و σ^2 تشخصان هذا التوزيع (احد اعضائه) وهما قيمتان مجهولتان عملياً الامر الذي يقتضي (ولاغراض التطبيقات الاحصائية) ايجاد تقدير عددي لكل منها . ان التقدير العددي للمعلمة (او المعالم) يمكن الحصول عليه على اساس قياسات عينة عشوائية قوامها n مسحوبة من مجتمع معرف بالدالة $f(x)$. هذا التقدير يسمى « المؤشر الاحصائي » . وهذا يعني ان المؤشر الاحصائي دالة بدلالة قياسات العينة خالية من اي مجهول . فاذا فرضنا ان θ تمثل معلمة وان $\hat{\theta}$ تقدير لهذه المعلمة فان $\hat{\theta} = g(x_1, x_2, \dots, x_n)$. فمثلاً اذا كانت $x_1, x_2, \dots, x_n$ قياسات عينة عشوائية من المفردات مسحوبة من $N(\mu, \sigma^2)$ فان كل من $\frac{x_1 + x_2}{2}$ ، $\sum_{i=1}^{n} \log x_i$ ، $\sum_{i=1}^{n} x_i^2$ يعتبر مؤشر احصائي في حين ان $x_i - \mu$ ، $\frac{\Sigma x_i}{\sigma}$ لايمكن اعتبارها مؤشرات احصائية بسبب اعتمادها على معالم مجهولة . ويمكن اعتبار $\hat{\theta}$ كافضل تقدير (من بين جملة تقديرات اخرى) الى θ اذا امتاز $\hat{\theta}$ بالخصائص التالية التي سنذكرها فقط دون اية تفاصيل كونها تقع في اختصاص الاستدلال الاحصائي statistical inference الذي من شأنه البحث عن ذلك التقدير الذي يتصف بكونه : غير متحيز unbiased ، متسق consistant ، كفوء efficient ، كافٍ sufficient ، ذو اقل تباين min . variance . لقد سبق وان ذكرنا في الفقرة (٨ ـ ١ ـ ١) ان عدد العينات الممكنة الاختيار من المجتمع هو C_n^N وهذا يعني ان قيمة $\hat{\theta}$ سوف تختلف من عينة لاخرى ، بحكم اختلاف مفردات هذه العينة كلياً او جزئياً ، الامر الذي يستدعي اعتبار θ أيضاً متغير عشوائي يسلك وفق دالة كثافة (او كتلة) احتمالية .

٨ ـ ١ ـ ٢ : توزيع متوسط العينة وتباينها

Distribution of sample mean and sample variance.

لتكن $x_1, x_2, \ldots, x_n$ قياسات عينة عشوائية قوامها n مفردة مسحوبة من (مجتمع توزيع) معرف بدالة كثافة احتمالية $f(x)$ (او كتلة احتمالية $p(x)$). وافرض ان σ^2, μ يمثلان على التوالي (اذا كانت موجودة) الوسط والتباين لهذا التوزيع . عندئذٍ :

$$\bar{x} = g_1(x_1, x_2, \ldots x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

يمثل الوسط الحسابي لقياسات هذه العينة وهو تقدير الى μ . ان $\bar{x}$ في ذات الوقت يعد متغيراً عشوائياً ذا توزيع معرف بالكثافة $f(x)$ التي تعتمد على $f(x)$. وعندئذٍ فان التوقع الى $\bar{X}$ هو :

$$E\bar{X} = \frac{1}{n} E \sum_{i=1}^{n} X_i = \frac{1}{n} \sum_{i=1}^{n} EX_i = \frac{1}{n} \sum_{i=1}^{n} \mu = \mu$$

وان تباين $\bar{X}$ هو :

$$V(\bar{X}) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} V(X_i)$$ (طالما ان X_i تمثل قياسات عينة عشوائية)

$$= \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}$$

وهذا يعني ان متوسط العينة $\bar{X}$ (التقدير الى μ) امتلك عزماً ذا مرتبة اولى حول نقطة الاصل هو μ وعزماً مركزياً ذا مرتبة ثانية هو $\frac{\sigma^2}{n}$ (بالاضافة الى امكانية تحديد عزوم اخرى من مراتب مختلفة الى $\bar{X}$) وذلك يعني ان $\bar{X}$ هو متغير عشوائي يسلك وفق دالة احتمالية مثل $f(\bar{x})$ او $p(\bar{x})$ بوسط قدره μ

وتباين $\frac{\sigma^2}{n}$ ان الدالة الاحتمالية الى $\bar{X}$ تعتمد بطبيعة الحال على دالة التوزيع الاحتمالي للمتغير العشوائي X الذي اختيرت منه تلك العينة .
كذلك فان ،

$$S^2 = g_2(x_1, x_2, \ldots, x_n) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

يمثل التباين للعينة وهو تقدير الى σ^2 وان S^2 في ذات الوقت متغير عشوائي يسلك وفق دالة احتمالية مثل $f(S^2)$ التي تعتمد ايضاً على دالة التوزيع الاحتمالي للمتغير X . ان توقع S^2 هو ،

$$ES^2 = \frac{1}{n-1} E\sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} E(X_i - \bar{X})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} E[(X_i - \mu) - (\bar{X} - \mu)]^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} E[E(X_i - \mu)^2 + E(\bar{X} - \mu)^2 - 2E(X_i - \mu)(\bar{X} - \mu)]$$

$$= \frac{1}{n-1} \sum_{i-1}^{n} \left[\sigma^2 + \frac{\sigma^2}{n} - \frac{2\sigma^2}{n} \right]$$

حيث ان

$$E(X_i - \mu)(\bar{X} - \mu) = \frac{1}{n} E(X_i - \mu) . \sum_{i=1}^{n} (X_i - \mu)$$

$$= \frac{1}{n} E[(X_i - \mu)(X_1 - \mu) + (X_i - \mu)(X_2 - \mu) + \ldots + (X_i - \mu)(X_n - \mu)]$$

$$= \frac{1}{n} E(X_i - \mu)^2$$

$$= \frac{\sigma^2}{n} \quad ; E(X_i - \mu)(X_j - \mu) = 0 \; \forall \; i \neq j$$

فاذن

$$ES^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left[\sigma^2 - \frac{\sigma^2}{n} \right] = \frac{1}{n-1} \left[n\sigma^2 - \sigma^2 \right]$$

$$\therefore ES^2 = \sigma^2$$

وان تباين S^2 هو :

$$V(S^2) = \frac{1}{(n-1)^2} V\left(\sum_{i=1}^{n} (X_i - \bar{X}^2) \right)$$

$$= \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4 \right) ; \mu_4 = E(X - \mu)^4$$

ونترك برهنة ذلك للقارىء . مما تقدم نلاحظ ان تباين العينة S^2 (التقدير الى σ^2) امتلك عزماً ذا مرتبة اولى حول نقطة الاصل هو σ^2 وعزماً مركزياً ذا مرتبة ثانية هو $V(S^2)$ وفق الصيغة اعلاه . وهذا يعني ان S^2 متغير عشوائي يسلك وفق دالة احتمالية مثل $f(S^2)$ بوسط قدره σ^2 وتباين مقداره $V(S^2)$.

٨ ـ ٢ : قانون الاعداد الكبيرة Law of large numbers

تطرقنا في الفقرة (٨ ـ ١) الى استعراض موجز لمفهوم العينة واساليب اختيارها وذكرنا ان الهدف الاساس من العينة هو حساب بعض المؤشرات الاحصائية كتقديرات لمعالم المجتمع الذي اختيرت منه تلك العينة . وكما هو معلوم فان حجم العينة يلعب دوراً اساسياً في دقة التقديرات التي نحصل عليها من تلك العينة . فكلما كان حجمها كبير فذلك يعني ان احتمال الفرق بين التقدير $\hat{\theta}$ والمعلمة θ

سوف يكون صغيراً . فعلى فرض ان $f(x)$ تمثل الدالة الاحتمالية الى X وان μ تمثل الوسط الى X في هذاالتوزيع . وافرض اننا نرغب في تقدير قيمة μ فان ذلك امر ممكن من خلال حساب $\bar{X}$ على اساس قياسات عينة مختارة من $f(x)$ قوامها n مفردة . الا انه وبشكل عام $\bar{X} \neq \mu$ لكن $E\bar{X} = \mu$ وان هدفنا الاساس هو جعل الفرق المطلق بين $\mu, \bar{X}$ ، اي $|\bar{X} - \mu|$ ، قريب من الصفر . ان لحجم العينة n دوراً اساسياً في تحقيق هذا الهدف من خلال ما يسمى بـ « قانون الاعداد الكبيرة » الذي ينص بما يلي : بفرض ان ε ، δ عددان صغيران بحيث ان $0 < \delta < 1$ ، $\varepsilon > 0$. عندئذٍ يوجد عدد موجب صحيح مثل n بحيث انه لو تم اختيار عينة عشوائية بحجم n او اكثر من توزيع معرف بالدالة $f(x)$ وتم حساب $\bar{X}$ فان احتمال ان يكون الفرق المطلق بين $\mu, \bar{X}$ اقل من ε هو اكبر من $1 - \delta$. وبالرموز فان :

$$P_r\{|\bar{X} - \mu|\varepsilon\} \geq 1 - \delta$$

ولغرض برهنة ذلك لابدلنا اولاً من اشتقاق ما يسمى بـ « متباينة تشيبيشيف » التي لها دور كبير في اشتقاق هذا القانون .

٨ ـ ٢ ـ ١ : متباينة تشيبيشيف Chebyshev's inequality

افرض ان X متغير عشوائي بدالة احتمالية $f(x)$ وبوسط وتباين محدودين هما على التوالي σ^2, μ . افرض ان k عدد موجب . عندئذٍ :

$$P_r[|X - \mu| < k\sigma] \geq 1 - \frac{1}{k^2} \quad \text{او ان} \quad P_r[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

البرهان : سوف نبرهن هذه المتباينة في حالة X من النوع المستمر ، والبرهان ذاته ينطبق في حالة X من النوع المتقطع وبمجرد استبدال رمز التكامل برمز الجمع .
من المعلوم ان :

$$\sigma^2 = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 . f(x)\,dx$$

$$= \int_{-\infty}^{\mu - k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x-\mu)^2 f(x)\,dx$$

$$+ \int_{\mu + k\sigma}^{\infty} (x-\mu)^2 . f(x)\,dx$$

$$\geq \int_{-\infty}^{\mu - k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu + k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx \quad \ldots(*)$$

واضح في التكامل الاول من(*) ان $x \leq \mu - k\sigma$ وذلك يعني ان $\mu - x \geq k\sigma$
وانه في التكامل الثاني من (*) $x \geq \mu + k\sigma$ وذلك يعني ان $x - \mu \geq k\sigma$
فاذن :

$$\sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} (k\sigma)^2 f(x)\,dx + \int_{\mu + k\sigma}^{\infty} (k\sigma)^2 f(x)\,dx$$

$$= k^2 \sigma^2 \left[\int_{-\infty}^{\mu - k\sigma} f(x)\,dx + \int_{\mu + k\sigma}^{\infty} f(x)\,dx \right.$$

$$= k^2 \sigma^2 [\, P_r(X \leq \mu - k\sigma) + P_r(X \geq \mu + k\sigma)\,]$$

$$= k^2 \sigma^2 [\, P_r(X - \mu \leq -k\sigma) + P_r(X - \mu \geq k\sigma)\,]$$

$$= k^2 \sigma^2 P_r[\,|X - \mu| \geq k\sigma]$$

$$\therefore \sigma^2 \geq k^2 \sigma^2 P_r[\,|X - \mu| \geq k\sigma\,] \quad \ldots(**)$$

وبقسمة طرفي (**) على $k^2 \sigma^2$ نحصل على :

$$P_r[\,|X - \mu| \geq k\sigma\,] \leq \frac{1}{k^2}$$

او ان :

$$1 - P_r[|X - \mu| \geq k\sigma] \geq 1 - \frac{1}{k^2}$$

اي ان

$$P_r[|X - \mu| < k\sigma] \geq 1 - \frac{1}{k^2}$$

وفرض ان $c = k\sigma > 0$ عندئذٍ فان هذه المتباينة ستكون :

$$P_r[|X - \mu| < c] \geq 1 - \frac{\sigma^2}{c^2} \text{ او ان } P_r[|X - \mu| \geq c] \leq \frac{\sigma^2}{c^2}$$

مثال (١) : افرض ان $X \sim N(40, 4)$. جد الحد الاعلى لاحتمال ان $|X - \mu| \geq 5$ ثم قارن هذا الاحتمال مع الاحتمال الحقيقي لحدوث الحادثة $\{|X - \mu| \geq 5\}$

الحل : واضح من معطيات السؤال ان $\mu = 40, \sigma = 2, c = 5$. فاذن :

$$P_r[|X - 40| \geq 5] \leq \frac{4}{25} = 0{\cdot}16$$

فاذن الحد الاعلى لاحتمال حدوث هذه الحادثة هو 0.16 . في حين ان الاحتمال الحقيقي لها هو :

$$\begin{aligned} P_r[|X - 40| \geq 5] &= 1 - P_r[|X - 40| < 5] \\ &= 1 - P_r(-5 < X - 40 < 5) \\ &= 1 - [P_r(X < 45) - P_r(X < 35)] \\ &= 1 - [P_r(Z < 2{\cdot}5) - P_r(Z < -2{\cdot}5)] \end{aligned}$$

من جداول التوزيع الطبيعي نجد ان

$$P_r(Z < -2{\cdot}5) = 0{\cdot}0062, \; P_r(Z < 2{\cdot}5) = 0{\cdot}9938$$

فاذن

$$P_r[|X - 40| \geq 5] = 0{\cdot}0124$$

مثال (٢) : افرض ان X متغير عشوائي بوسط μ وتباين $\sigma^2 = 0$. برهن ان $P_r(X = \mu) = 1$

الحل : ان $\sigma^2 = 0$ ، ولاي عدد موجب مثل c لدينا :

$$P_r[\,|X - \mu| \geq c\,] \leq 0$$

وحيث ان الاحتمال قيمة غير سالبة لذا :

$$P_r[\,|X - \mu| \geq c\,] = 0 \quad \forall c$$

وذلك يعني ان احتمال حدوث فرق مطلق بين μ, X مساوٍ للصفر ، اي انه يجب ان يكون $P_r(X = \mu) = 1$

٨ ـ ٢ ـ ٢ : برهان قانون الاعداد الكبيرة

افرض ان $X_1, X_2, \ldots, X_n$ متغيرات عشوائية مستقلة (او عينة عشوائية) تتوزع وفق نفس دالة الكثافة الاحتمالية $f(x)$ (او كتلة احتمالية $P(x)$) . وافرض ان $V(X_i) = \sigma^2, EX_i = \mu$. ليكن $\bar{X}$ يمثل الوسط الحسابي لهذه المتغيرات وان $v(\bar{X}) = \sigma^2 / n, E\bar{X} = \mu$. عندئذٍ

$$P_r[\,|\bar{X} - \mu| \geq c\,] \leq \frac{\sigma^2}{nc^2}, \quad c > 0$$

وان $\bar{X}$ يقترب من μ عندما $n \to \infty$

البرهان : حسب متباينة تشيبيشيف فان

$$P_r[\,|\bar{X} - E\bar{X}| \geq c\,] \leq \frac{V(\bar{X})}{c^2}$$

لكن $E\bar{X} = \mu$, $V(\bar{X}) = \dfrac{\sigma^2}{n}$ ، فاذن

$$P_r[\,|\bar{X} - \mu| \geq c\,] \leq \frac{\sigma^2}{nc^2}$$

وبفرض ان $n \to \infty$ عندئذٍ

$$\lim_{n \to \infty} P_r[|\bar{X} - \mu| \geq c] \leq \lim_{n \to \infty} \frac{\sigma^2}{nc^2} = 0$$

فاذن

$$\lim_{n \to \infty} P_r[|\bar{X} - \mu| \geq c] = 0$$

وهذا يعني ان احتمال الفرق المطلق بين $\bar{X}, \mu$ هو اكبر من او يساوي c مساوٍ للصفر عندما $n \to \infty$ ، او ان

$$\lim_{n \to \infty} P_r[|\bar{X} - \mu| < c] = 1$$

وهذا يعني انه باحتمال قدرة واحد $\bar{X}$ يقترب من μ عندما $n \to \infty$.
وفق ماتقدم نقول ان $\bar{X}$ يتقارب بالاحتمال من μ. ويرمز لذلك بالشكل: $\bar{X} \xrightarrow{P} \mu$.

مثال (٣) : باحتمال لايقل عن 0.95 ، جد حجم العينة n المطلوب سحبها من مجتمع احصائي الذي يجعل الفرق المطلق بين $\bar{X}, \mu$ لايزيد عن $\frac{\sigma}{10}$.

الحل : ان $c = \frac{\sigma}{10}$. عندئذٍ

$$P_r\left[|\bar{X} - \mu| < \frac{\sigma}{10}\right] \geq 1 - \frac{\sigma^2}{n\left(\frac{\sigma}{10}\right)^2} = 1 - \frac{100}{n}$$

وحيث ان الاحتمال المعطى لايقل عن 0·95 فذلك يعني ان

$$0{\cdot}95 = 1 - \frac{100}{n} \to \frac{100}{n} = 0{\cdot}05 \to n = 2000$$

وهذا يعني ان حجم العينة المطلوب سحبها يجب ان لايقل عن 2000 مفردة وفق هذه المعطيات.

مثال (٤٠) : افرض ان $X \sim N(\mu, 4)$. جد حجم العينة n المطلوب سحبها من هذا المجتمع التي تجعل احتمال الفرق المطلق بين $\bar{X}, \mu$ اقل من 0.9 لايقل عن 0.99

الحل : ان $c = 0.9$. فاذن

$$P_r [|\bar{X} - \mu| < 0.9] \geq 1 - \frac{4}{n(0.9)^2} = 0.99$$

وهذا يعني ان

$$\frac{4}{n(0.9)^2} = 0.01 \rightarrow n = \frac{4}{0.0081} \simeq 494$$

عليه فان حجم العينة المطلوب وفق معطيات هذا المثال يجب ان لايقل عن 494 مفردة .

٨ ـ ٢ ـ ٣ : مبرهنة الغاية المركزية Central limit theorem

تعتبر مبرهنة الغاية المركزية احدى الركائز الاساسية في النظرية الاحصائية وذات فائدة تطبيقية كبيرة وخصوصاً مايتعلق الامر بحساب الاحتمالات وموضوع اختبار الفرضيات الاحصائية . لقد سبق ذكر هذه المبرهنة في فقرات عديدة من الفصول السابقة وبشكل غير مباشر . وفيما يلي نص وبرهان هذه المبرهنة علماً ان هنالك اشكال اخرى لبرهنتها كل منها يفترض فروض وشروط معينة الا ان الهدف من البرهان هو نفسه .

افرض ان X متغير عشوائي بدالة كثافة احتمالية $f(x)$ او كتلة احتمالية $P(x)$. وافرض ان $EX = \mu$, $V(X) = \sigma^2 < \infty$. ليكن $\bar{X}$ يمثل الوسط الحسابي لقياسات عينة عشوائية قوامها n مسحوبة من مجتمع الدالة $f(x)$ او $P(x)$ ، وان $Y = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$. عندئذٍ فان توزيع Y يقترب من التوزيع الطبيعي المعياري عندما $n \rightarrow \infty$.

البرهان: ليكن $X^* \sim N(\mu^*, \sigma^{*2})$. عندئذٍ

$Y^* = \dfrac{X^* - \mu^*}{\sigma^*} \sim N(0,1)$ وان $M_{Y^*}(t) = e^{\frac{1}{2}t^2} = M_1(t)$.

افرض ان $M_X(t)$ موجودة ومستمرة لجميع قيم $-h < t < h$، $h > 0$ عندئذٍ فان الدالة المولدة لعزوم $(X - \mu)/\sigma$ موجودة وهي $M_2(t) = Ee^{t\left(\frac{X-\mu}{\sigma}\right)}$

وبفرض ان $\bar{X}$ يمثل الوسط الحسابي لقياسات عينة عشوائية مسحوبة من $f(x)$ او $P(x)$ قوامها n مفردة ، وان $\bar{X}$ يسلك وفق دالة احتمالية معينة مثل $g(\bar{X})$ عندئذٍ فان $E\bar{X} = \mu$ ، $V(\bar{X}) = \dfrac{\sigma^2}{n}$.

افرض ان $Z = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}}$. عندئذٍ فان الدالة المولدة لعزوم Z هي

$$M_Z(t) = Ee^{tz} = Ee^{t\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)} = M_3(t)$$

هدفنا الان هو البرهنة على ان

$$\lim_{n \to \infty} M_3(t) = M_1(t)$$

ان

$$M_3(t) = Ee^{\frac{t}{n}\left(\frac{n\bar{X} - n\mu}{\sigma/\sqrt{n}}\right)} = Ee^{\frac{t}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)}$$

وحيث ان $X_1, X_2, \ldots, X_n$ تشكل قياسات عينة عشوائية فهي بحكم المتغيرات العشوائية المستقلة ذات نفس التوزيع . وهذا يعني ان

$$M_3(t) = \prod_{i=1}^{n} Ee^{\frac{t}{\sqrt{n}}\left(\frac{X_i - \mu}{\sigma}\right)} = \prod_{i=1}^{n} M_2\left(\frac{1}{\sqrt{n}}\right)$$

فاذن

$$M_3(t) = \left[M_2\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

لكن

$$M_2(t) = Ee^{t\left(\frac{X-\mu}{\sigma}\right)} = E\sum_{j=0}^{\infty}\frac{(tv)^j}{j!}, v = \frac{X-\mu}{\sigma}$$

$$= E\left[1 + tv + \frac{t^2}{2!}.v^2 + \frac{t^3}{3!}v^3 + \ldots + \frac{t^r}{r!}v^r + \ldots\right]$$

$$= 1 + \frac{t^2}{2!}Ev^2 + \frac{t^3}{3!}Ev^3 + \ldots + \frac{t^r}{r!}Ev^r + \ldots\Bigg], Ev = 0$$

عليه فان

$$M_2\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2!n} + \frac{t^3}{3!n\sqrt{n}}.Ev^3 + \ldots + \frac{t^r}{r!.n^{r/2}}Ev^r + \ldots, Ev^2 = 1$$

$$= 1 + \frac{1}{n}\left[\frac{t^2}{2} + \frac{t^3}{6\sqrt{n}}EV^3 + \ldots + \frac{t^r}{r!n^{r/2-1}}EV^r + \ldots\right]$$

$$= 1 + \frac{1}{n}K$$

فاذن

$$M_3(t) = \left[1 + \frac{K}{n}\right]^n$$

وان

$$\lim_{n\to\infty}\left[1 + \frac{K}{n}\right]^n = e^{\lim_{n\to\infty} K}$$

حيث ان K دالة في n .

$$= e^{\lim_{n\to\infty}\left(\frac{t^2}{2} + \frac{t^3}{6\sqrt{n}}EV^3 + \ldots + \frac{t^r}{r!n^{r/2-1}} + \ldots\right)}$$

فاذن

$$\lim_{n\to\infty} M_3(t) = e^{\frac{1}{2}t^2}$$

لكن $e^{\frac{1}{2}t^2}$ تمثل الدالة المولدة لعزوم توزيع طبيعي معياري . فاذن نستنتج ان

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \xrightarrow{n \to \infty} N(0, 1)$$

وبشكل عام نستنتج من هذه المبرهنة ان الدرجة المعيارية في اي توزيع احتمالي معرف بالدالة $f(x)$ او $P(x)$ يؤول توزيعها الى التوزيع الطبيعي المعياري عندما $n \to \infty$ فمثلاً اذا كان $X \sim b(n, P)$ فان

$$Z = \frac{X - np}{\sqrt{npq}} \xrightarrow{n \to \infty} N(0, 1)$$

او ان

$$X \sim N(np, npq), n \to \infty$$

كذلك اذا كان $X \sim P0(m)$ فان :

$Z = \frac{X - m}{\sqrt{m}} \xrightarrow{m \to \infty} N(0, 1)$ او ان $X \sim N(m, m)$ ، وهكذا الحال بالنسبة لاي توزيع احتمالي اخر بشرط ان $\sigma^2 < \infty$. وبفرض ان $X \sim b(n, P)$ وتطلب الامر حساب الاحتمال $P_r(X \leq x_0)$, x_0 قيمة معطاة ، عندما n تكون كبيرة فان ذلك يتم وفق الآتي :

$$P_r(X \leq x_0) = P_r\left(\frac{X - np}{\sqrt{npq}} \leq \frac{x_0 - np}{\sqrt{npq}} \right)$$

$$= P_r(Z \leq z_0), Z \sim N(0, 1)$$

وهكذا الامر من حيث المفهوم عند حساب احتمال معين لمتغير عشوائي يسلك وفق دالة كثافة احتمالية او دالة كتلة احتمالية . اي انه وبشكل عام :

$$P_r(X \leq x_0) = P_r\left(\frac{X - EX}{\sqrt{V(X)}} \leq \frac{x_0 - EX}{\sqrt{V(X)}} \right)$$

$$= P_r(Z \leq z_0), Z \sim N(0, 1), n \to \infty$$

وغالبا مايتم اضافة 0.5 الى المقدار $(x_0 - EX)$ حيث ان هذا المقدار يسمى بـ « مصحح الاستمرارية Continuity Correction » يضفي الى قيمة الاحتمال المستخرج دقة اكبر . ومن الناحية التطبيقية فان درجة التقارب من التوزيع الطبيعي تعتمد بطبيعة الحال على n وعلى دالة التوزيع الاحتمالي للمتغير X .

مثال (٥) : افرض ان $X \sim b(1000, 0.4)$. جد $P_r(X \leq 420)$

الحل :

ان مسألة ايجاد $P_r(X \leq 420)$ باستخدام توزيع ثنائي الحدين تبدو معقدة نظراً لان ذلك يتطلب حساب :

$$P_r(X \leq 420) = \sum_{x=0}^{420} C_x^{1000} (0.4)^x . (0.6)^{1000-x}$$

لكن ، وحيث ان n كبيرة فانه يمكن حساب قيمة تقريبية لهذا الاحتمال وفق الآتي :

$$npq = (1000)(0.4)(0.6), np = (1000)(0.4) = 400$$

$$= 240 \quad \therefore \sqrt{npq} = 15.492$$

فاذن

$$P_r(X \leq 420) = P_r\left(Z \leq \frac{420 - 400}{15.492}\right)$$

$$= P_r(Z \leq 1.29)$$

من جداول التوزيع الطبيعي نجد ان قيمة هذا الاحتمال هي تقريباً 0.9015 في حين لو تم استخدام مصحح الاستمرارية فان :

$$P_r(X \leq 420) = P_r\left(Z \leq \frac{420 + 0{\cdot}5 - 400}{15{\cdot}492}\right)$$

$$= P_r(Z \leq 1{\cdot}32) \simeq 0{\cdot}9066$$

مثال (٦) : اذا علمت ان $X \sim P0(36)$ جد $P_r(X \geq 40)$.

الحل :

$$P_r(X \geq 40) = P_r\left(Z \geq \frac{40 - 36}{6}\right) = P_r(Z \geq 0{\cdot}67)$$

$$= 1 - P_r(Z < 0{\cdot}67)$$

من جداول التوزيع الطبيعي نجد ان

$$P_r(Z < 0{\cdot}67) = 0{\cdot}7486$$

فاذن

$$P_r(X \geq 40) \simeq 0{\cdot}2514$$

من خلال هذين المثالين يمكن الاستنتاج بما يلي :

اذا كانت $X_1, X_2, \dots, X_n$ متغيرات عشوائية مستقلة تتوزع وفق نفس دالة الكثافة الاحتمالية $f(x)$ او دالة كتلة احتمالية $P(x)$ بوسط μ وتباين مُحدود σ^2 . وان $\bar{X}_n$ يمثل الوسط الحسابي لهذه المتغيرات $Z_n = (\bar{X}_n - E\bar{X}_n) / \sqrt{V(\bar{X}_n)}$ وبفرض ان $F_{Z_n}(z)$ تمثل الدالة التوزيعية الى Z_n . عندئذٍ واستناداً لمبرهنة الغاية المركزية فان $F_{Z_n}(z)$ تتقارب من $\Phi(z)$ عندما $n \to \infty$ ، حيث ان $\Phi(\cdot)$ تعني الدالة التوزيعية في التوزيع الطبيعي المعياري . اي ان :

$$\lim_{n \to \infty} F_{Z_n}(z) = \Phi(z)$$

وهذا يعني انه لاي عددين حقيقيين مثل $a < b, b, a$ فان :

$$P_r(a \leq \bar{X}_n \leq b) = P_r\left(\frac{a - \mu}{\sigma / \sqrt{n}} \leq Z_n \leq \frac{b - \mu}{\sigma / \sqrt{n}}\right)$$

$$= P_r(Z_1 \leq Z_n \leq Z_2) = \Phi(Z_2) - \Phi(Z_1), n \to \infty$$

٨ ـ ٢ ـ ٤ : التوزيعات المقيدة والتقارب التصادفي Limiting distributions and stochastic convergence

لاحظنا في فصول وفقرات سابقة ان بعض التوزيعات الاحتمالية لمتغيرات عشوائية (او لمؤشرات احصائية كالوسط الحسابي $\bar{X}$ مثلاً) كانت تعتمد على حجم العينة n. فمثلاً لاحظنا ان التوزيع الاحتمالي الى $\bar{X}$ لعينة عشوائية مسحوبة من $N(\mu, \sigma^2)$ كان $N(\mu . \sigma^2/n)$ ، وذلك يعني ان $f(\bar{x})$ تعتمد على n وان الدالة المولدة لعزوم $\bar{X}$ هي الاخرى سوف تعتمد على n وكذلك الدالة التوزيعية $F(\bar{x})$. وفي بعض الاحيان ولدى استخدامنا للدالة المولدة للعزوم كاسلوب لاستنتاج التوزيع الاحتمالي لمتغير (أو مؤشر احصائي) كدالة بدلالة متغير آخر (او متغيرات اخرى) قد نقف امام حالة تكون فيها مسألة استنتاج التوزيع الاحتمالي لذلك المتغير (او المؤشر) صعبة او غير ممكنة مما يضطرنا الامر الى عمل بعض التقريبات التي يمكن من خلالها التوصل لتلك الدالة عند توفر شروط معينة . فمثلاً اذا كان X يتوزع وفق دالة التوزيع المنتظم على الفترة (0 , 1) وان $\bar{X}$ يمثل الوسط الحسابي لعينة مسحوبة من هذا التوزيع ، فان :

$$M_{\bar{X}}(t) = Ee^{t\bar{X}} = Ee^{\frac{t}{n}\sum_{i=1}^{n} X_i}$$

$$= \prod_{i=1}^{n} M_{X_i}\left(\frac{t}{n}\right)$$

لكن

$$M_{X_i}(t) = \frac{e^t - 1}{t} \quad \forall_{X_i}, t \neq 0$$

فاذن

$$M_{\bar{X}}(t) = \left[\frac{e^{t/n} - 1}{t/n}\right]^n$$

لاحظ ان الدالة المولدة لعزوم $\bar{X}$ تعتمد على n وهذا يعني ان التوزيع الاحتمالي الى $\bar{X}$ هو الاخر يعتمد على n الذي يفترض انه موجود طالما ان $M_{\bar{X}}(t)$ موجودة (حسب صفة الوحدانية للدوال المولدة للعزوم) الا ان مسألة التوصل لتوزيع $\bar{X}$ صعبة وفق الاساليب التي درسناها في الفصل السابع مما يتطلب

الامر البحث عن هذا التوزيع وفق شروط مفروضة على n . توزيعات من هذا النوع تسمى « توزيعات مقيدة » . وفيما يلي تعريف لهذا النوع من التوزيعات .

افرض ان $F_n(y)$ تمثل الدالة التوزيعية للمتغير العشوائي (او المؤشر الاحصائي) Y_n التي تعتمد على n ، وافرض ان F (y) تمثل الدالة التوزيعية الى Y. فاذا كان $\lim_{n \to \infty} F_n(y) = F(y)$ لاية قيمة مخصصة الى Y مثل y بحيث ان الدالة F (y) عند y تكون مستمرة وقابلة للاشتقاق عندئذ يقال ان المتغير العشوائي (او المؤشر الاحصائي) Y_n يمتلك توزيع مقيد وان الدالة التوزيعية له هي F (y)

مثال (٧) : افرض ان $\bar{X}$ يمثل الوسط الحسابي لقياسات عينة عشوائية قوامها n مسحوبة من N (0, 1) . عندئذٍ فان $\bar{X} \sim N\left(0, \frac{1}{n}\right)$ وان

$$F_n(\bar{x}) = \int_{-\infty}^{\bar{x}} \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{1}{2}n\bar{x}^2} d\bar{x} \; ; -\infty < \bar{x} < \infty$$

وبفرض ان

$$Z = \sqrt{n}\,\bar{x} \rightarrow \bar{x} = \frac{Z}{\sqrt{n}} \rightarrow d\bar{x} = \frac{dz}{\sqrt{n}}$$

فاذن

$$F_n(\bar{x}) = \int_{-\infty}^{\sqrt{n}\,\bar{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

ويتضح من الصيغة الاخيرة ان :

$$\begin{aligned}\lim_{n \to \infty} F_n(\bar{x}) &= 0 \quad, \bar{x} < 0 \\ &= 0{\cdot}5 \,, \bar{x} = 0 \\ &= 1 \quad, \bar{x} > 0\end{aligned}$$

وان الدالة :

$$\begin{aligned}F(\bar{x}) &= 0 \quad, \bar{x} < 0 \\ &= 1 \quad, \bar{x} \geq 0\end{aligned}$$

تمثل دالة توزيعية وان $\lim_{n \to \infty} F_n(\bar{x}) = F(\bar{x})$ عند اية نقطة استمرارية الى $F(\bar{x})$ عدا النقطة $\bar{x} = 0$ التي عندها يكون $\lim F_n(0) \neq F(0)$ عليه فان المتغير العشوائي $\bar{X}_n$ يمتلك توزيع مقيد بدالة توزيعية $(\bar{x})$ لكنه توزيع غير متولد (اي لم نحصل على دالة احتمالية بدلالة $\bar{x}$) وان كل الاحتمال الممكن قد اقترن بنقطة واحدة فقط هي $\bar{x} = 0$.

مثال (٨) : افرض ان $X_1, X_2, \ldots, X_n$ تمثل عينة عشوائية من توزيع منتظم مستمر على الفترة $(0, \theta)$ ، $\theta > 0$ ، وبفرض ان $Y_n = \text{Max}(X_i)$ ، $i = 1, 2, \ldots, n$ وان دالة الكثافة الاحتمالية الى Y_n هي $0 < y < \theta$ ، $g_n(y) = \frac{ny^{n-1}}{\theta^n}$. وافرض ان $z_n = n(\theta - Y_n)$. عندئذٍ وباستخدام اسلوب التحويلات يمكن ملاحظة ان دالة الكثافة الاحتمالية الى z_n هي

$$h_n(Z) = \frac{(\theta - Z/n)^{n-1}}{\theta^n}, 0 < z < n\theta$$

ان الدالة التوزيعية للمتغير z_n هي :

$$G_n(z) = \int_0^z \frac{(\theta - z/n)^{n-1}}{\theta^n} dz$$

$$= -\frac{n}{\theta^n}\left[\frac{(\theta - z/n)^n}{n}\right]_0^z$$

$$= 1 - \left(1 - \frac{z}{n\theta}\right)^n ; 0 < z < n\theta$$

ويتضح من هذه الدالة ان :

$$G_n(z) = 0 \quad , z \leq 0$$

$$= 1 - \left(1 - \frac{z}{n\theta}\right)^n , 0 < z < n\theta$$

$$= 1 \quad , z \geq n\theta$$

عليه فان :

$$\lim_{n \to \infty} G_n(z) = \lim_{n \to \infty} \left[1 - \left(1 - \frac{z}{n\theta} \right)^n \right]$$

$$= \lim_{n \to \infty} \left[1 - \left(1 - \frac{z/\theta}{n} \right)^n \right]$$

$$= 1 - \lim_{n \to \infty} \left(1 - \frac{z/\theta}{n} \right)^n$$

$$= 1 - e^{-\frac{z}{\theta}}, z \geq 0$$

$$= G(z)$$

ويتضح ان :

$$\lim_{n \to \infty} G_n(z) = 0 \qquad , z \leq 0$$

$$= 1 - e^{-\frac{z}{\theta}} , z > 0$$

وان

$$G(z) = 0 \qquad , z \leq 0$$

$$= 1 - e^{-\frac{z}{\theta}} , z > 0$$

وهذا يعني ان $G(z)$ مستمرة دائماً وان $\lim_{n \to \infty} G_n(z) = G(z)$. عليه فان z_n يمتلك توزيع مقيد بدالة توزيعية $G(z)$ وبذلك فان التوزيع المقيد الى z_n هو

$$g(z) = \frac{dG(z)}{dz} = \frac{1}{\theta} e^{-\frac{z}{\theta}} , z \geq 0$$

وهذا يعني ان التوزيع المقيد الى z_n هو توزيع اسي بالمعلمة $\frac{1}{\theta}$ ، اي انه توزيع متولد . مع ملاحظة ان التوزيع المقيد لمتغير عشوائي (او مؤشر احصائي) ليس بالضرورة ان يكون موجود دائماً . فقد نحصل على توزيع مقيد او قد لانحصل عليه .

واذا كانت $Y_1, Y_2, \ldots, Y_n, \ldots$ تمثل سلسلة من المتغيرات العشوائية ، عندئذٍ يقال ان للمتغير Y_n ، مثلاً ، توزيع محاذي Asymptotic distribution اذا كان توزيع Y_n مساو تقريباً للتوزيع الحقيقي الى Y_n عندما $n \to \infty$. فمثلاً اذا كان $\bar{X}_n$ يمثل الوسط الحسابي لعينة عشوائية مسحوبة من مجتمع معرف بالدالة $f(x)$ فان $\bar{X}_n$ يمتلك توزيعاً محاذياً هو توزيع طبيعي بوسط μ وتباين σ^2 / n عندما n تكون كبيرة . ان التوزيع المحاذي قد يعتمد على حجم العينة n في حين ان التوزيع المقيد لا يعتمد اطلاقاً على n بسبب اختزالها عند اخذ الغاية lim . .

اما مفهوم « التقارب التصادفي » فله علاقة بفهوم التوزيعات المقيدة . وسوف نكتفي بعرض تعريف وخصائص هذا المفهوم .

افرض ان $F_n(y)$ تمثل الدالة التوزيعية للمتغير العشوائي (او المؤشر الاحصائي) Y_n الذي يفترض ان توزيعه الاحتمالي يعتمد على حجم العينة n . ليكن c ثابت لا يعتمد على n . عندئذٍ يقال ان Y_n متقارب تصادفياً من الثابت c اذا وفقط اذا كان :

$$\lim_{n \to \infty} P_r[\,|Y_n - c| < \varepsilon\,] = 1 \quad \forall\, \varepsilon > 0$$

وان

$$\lim_{n \to \infty} P_r[\,|Y_n - c| \geq \varepsilon\,] = 0 \quad \forall\, \varepsilon > 0$$

مما تقدم نستنتج انه عندما يكون التوزيع المقيد لمتغير عشوائي مثل Y_n غير متولد عندئذٍ يقال ان Y_n متقارب تصادفياً من الثابت c الذي يمتلك احتمالاً قدره واحد . ففي المثال (٧) من هذه الفقرة لاحظنا ان $\bar{X}_n$ تقارب تصادفياً الى الصفر (وهو متوسط التوزيع الطبيعي الذي سحبت منه العينة) بحيث ان الاحتمال المقترن بهذه النقطة كان واحداً .

مثال (٩) : افرض ان $\bar{X}_n$ يمثل الوسط الحسابي لعينة عشوائية مسحوبة من $N(\mu, \sigma^2)$. عندئذٍ فان $X_n \sim N(\mu, \sigma^2/n)$. الان بفرض ان $0 < \varepsilon$ ، عندئذٍ :

$$P_r[\,|\bar{X}_n - \mu| \geq \varepsilon\,] = P_r\left[\,|\bar{X}_n - \mu| \geq \frac{k\sigma}{\sqrt{n}}\right]; k = \frac{\varepsilon\sqrt{n}}{\sigma}$$

وحسب متباينة تشيبيشيف فان هذا الاحتمال اقل من او يساوي $1/k^2$. فاذن

$$P_r\left[\,|\bar{X}_n - \mu| \geq \frac{k\sigma}{\sqrt{n}}\right] \leq \frac{1}{k^2} = \frac{\sigma^2}{n\varepsilon^2}$$

عليه فان :

$$\lim_{n\to\infty} P_r\left[\,|\bar{X}_n - \mu| \geq \frac{k\sigma}{\sqrt{n}}\right] \leq \lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$$

وهذا يعني ان :

$$\lim_{n\to\infty} P_r\left[\,|\bar{X}_n - \mu| < \frac{k\sigma}{\sqrt{n}}\right] = 1$$

اي ان $\bar{X}_n$ متقارب تصادفياً من μ .

مثال (١٠) : افرض ان $Y_n \sim b(n, P)$ بين ان $\frac{Y_n}{n}$ (التي تمثل نسبة عدد حالات النجاح) متقاربة تصادفياً من P .

الحل : افرض ان $0 < \varepsilon$. اذن

$$P_r\left[\left|\frac{Y_n}{n} - E\left(\frac{Y_n}{n}\right)\right| \geq \varepsilon\right] = P_r\left[\left|\frac{Y_n}{n} - P\right| \geq \varepsilon\right]; EY_n = nP.$$

وحسب متباينة تشيبيشيف فان هذا الاحتمال اقل من او يساوي

$$\frac{V\left(\frac{Y_n}{n}\right)}{\varepsilon^2} = \frac{Pq}{n\varepsilon^2}, q = 1 - P, V\left(\frac{Y_n}{n}\right) = \frac{Pq}{n}$$

فاذن

$$P_r\left[\left|\frac{Y_n}{n} - P\right| \geq \varepsilon\right] \leq \frac{Pq}{n\varepsilon^2}$$

عليه فان

$$\lim_{n \to \infty} P_r\left[\left|\frac{Y_n}{n} - P\right| \geq \varepsilon\right] \leq \lim_{n \to \infty} \frac{Pq}{n\varepsilon^2} = 0$$

وهذا يعني ان Y_n / n متقاربة تصادفياً من P .
ان مفهوم « التقارب التصادفي » يسمى في بعض الاحيان « التقارب بالاحتمال » Convergence in Probability . وبالرموز فان :

$$Y_n \xrightarrow[n \to \infty]{P} c$$

اي ان Y_n يتقارب بالاحتمال من c عندما $n \to \infty$. ومن اهم خصائص التقارب بالاحتمال مايلي :

اذا كان $X_n \xrightarrow[n \to \infty]{P} a, Y_n \xrightarrow[n \to \infty]{P} b$ عندئذٍ :

أ ـ ان

$$X_n \pm Y_n \xrightarrow[n \to \infty]{P} a \pm b$$

ب ـ ان

$$X_n \cdot Y_n \xrightarrow[n \to \infty]{P} a.b$$

جـ ـ ان

$$\frac{X_n}{Y_n} \xrightarrow[n \to \infty]{P} \frac{a}{b} \; ; b \neq 0$$

٨ ـ ٢ ـ ٥ : دوال توليد العزوم المقيدة

Limiting moment generating functions

سبق وان ذكرنا في الفصل الثاني لدى دراستنا لموضوع الدوال المولدة للعزوم بانه اذا كانت الدالة المولدة لعزوم متغير عشوائي موجودة فهي دالة وحيدة وتشخص التوزيع الاحتمالي لذلك المتغير. وفق هذا المفهوم فانه يمكن ايضاً استنتاج التوزيع الاحتمالي المقيد لمتغير عشوائي اذا علمت الدالة المولدة لعزومه المقيدة (اذا كانت موجودة).

افرض ان Y_n متغير عشوائي (أو مؤشر احصائي) بدالة توزيعية $F_n(y)$ تعتمد على n وبدالة مولده للعزوم مثل $M(t;n)$، التي تعتمد على n ايضاً، موجودة لجميع قيم t المعرفة بالفترة $(-h, h)$، $h > 0$. عندئذٍ اذا كانت هنالك دالة توزيعية مثل $F(y)$ ودالة مولده للعزوم مثل $M(t)$ للمتغير Y معرفة عند قيم $|t| \leq h_1 < h$ بحيث ان $\lim\limits_{n \to \infty} M(t;n) = M(t)$ عندئذٍ فان Y_n سوف يمتلك توزيع مقيد بدالة توزيعية $F(y)$.

وبهدف توضيح هذا الموضوع تأمل الغاية التالية التي يرد ذكرها في بعض مصادر الرياضيات المتقدمة :

$$\lim_{n \to \infty} \left[1 + \frac{a}{n} + \frac{g(n)}{n} \right]^{bn}$$

حيث ان a, b كميات لاتعتمد على n وان $g(n)$ دالة في n بحيث ان $\lim\limits_{n \to \infty} g(n) = 0$. عندئذٍ :

$$\lim_{n\to\infty}\left[1+\frac{a}{n}+\frac{g(n)}{n}\right]^{bn}=\lim_{n\to\infty}\left[1+\frac{a}{n}\right]^{bn}$$

$$=\lim_{n\to\infty}\left[1+\frac{ba}{bn}\right]^{bn}=\lim_{m\to\infty}\left[1+\frac{ba}{m}\right]^{m};m=bn$$

$$=e^{ba}$$

مثال (١١) : افرض ان الدالة المولدة لعزوم متغير عشوائي Y_n هي $M(t,n)=\left[1+\frac{t^2}{n}+\frac{t^3}{n\sqrt{n}}\right]^{\frac{n}{2}}$. جد التوزيع المقيد الى Y_n

الحل : بالمقارنة مع الغاية اعلاه نجد ان :

$$a=t^2,b=\frac{1}{2},g(n)=\frac{t^3}{\sqrt{n}},\quad \lim_{n\to\infty}g(n)=0$$

فاذن

$$\lim_{n\to\infty}M(t,n)=\lim_{n\to\infty}\left[1+\frac{t^2}{n}+\frac{t^3/\sqrt{n}}{n}\right]^{\frac{n}{2}}$$

$$=\lim_{n\to\infty}\left[1+\frac{t^2}{n}\right]^{\frac{n}{2}}$$

$$=\lim_{n\to\infty}\left[1+\frac{t^2/2}{n/2}\right]^{\frac{n}{2}}$$

$$=\lim_{m\to\infty}\left[1+\frac{t^2/2}{m}\right]^{m}=e^{t^2/2}=M(t)$$

لاحظ ان $M(t)$ هنا تمثل الدالة المولدة لعزوم $N(0,1)$ وهذا يعني ان التوزيع المقيد الى Y_n هو $N(0,1)$.

مثال (١٢) : افرض ان $Y_n \sim b(n,p)$ وان $\lambda = np$ حيث ان λ قيمة ثابتة لاية قيمة مخصصة الى n . جد التوزيع المقيد الى Y_n .

الحل : حيث ان $Y_n \sim b(n,p)$ فاذن

$$M(t,n) = (q + pe^t)^n = (1 - p + pe^t))^n$$

لكن $\lambda = np$. فاذن $p = \frac{\lambda}{n}$ عليه فان

$$M(t,n) = \left(1 - \frac{\lambda}{n} + \frac{\lambda e^t}{n}\right)^n$$

$$= \left(1 + \frac{\lambda}{n}(e^t - 1)\right)^n$$

$$= \left(1 + \frac{a}{n}\right)^n \;; a = \lambda(e^t - 1)$$

$$\therefore \lim_{n\to\infty} M(t,n) = \lim_{n\to\infty}\left(1 + \frac{a}{n}\right)^n = e^a$$

$$= e^{\lambda(e^t - 1)} = M(t)$$

لاحظ هنا ان $M(t)$ تمثل الدالة المولدة لعزوم توزيع پواسون بالمعلمة λ . فاذن نستنتج ان التوزيع المقيد الى Y_n هو توزيع پواسون بالمعلمة λ . وهذه تمثل طريقة اخرى لبيان ان توزيع پواسون هو توزيع تقاربي من توزيع ثنائي الحدين .

مثال (١٣) : افرض ان Y_n يتوزع كتوزيع ثنائي الحدين السالب بالمعلمتين p , n بحيث ان p قريبة جداً من الواحد . وان $\frac{nq}{p} = \lambda$ قيمة ثابتة لاية قيمة مخصصة الى n . جد التوزيع المقيد الى Y_n .

الحل : حيث ان $Y_n \sim Nb(n, p)$ فان :

$$M(t, n) = \left(\frac{p}{1 - qe^t} \right)^n \quad ; q = 1 - P, qe^t < 1$$

$$\therefore K(t, n) = \log \left(\frac{p}{1 - qe^t} \right)^n = n[\log p - \log(1 - qe^t)]$$

وبشكل عام وحسب متسلسلة تايلر فان :

$$\log(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots - \frac{x^m}{m} - \dots, |x| < 1$$

واستناداً لهذه المتسلسلة فان :

$$\log p = \log(1 - q) = -q - \frac{q^2}{2} - \frac{q^3}{3} - \dots - \frac{q^m}{m} - \dots$$

وان

$$\log(1 - qe^t) = -qe^t - \frac{1}{2} q^2 e^{2t} - \frac{1}{3} q^3 e^{3t} - \dots$$

فاذن

$$K(t, n) = n \left[q(e^t - 1) + \frac{1}{2} q^2 (e^{2t} - 1) + \dots \right] \dots (*)$$

وحيث ان $\lambda = \frac{nq}{p}$ فاذن $q = \frac{\lambda p}{n}$ وبالتعويض عن q في (*) نحصل على :

$$K(t, n) = \lambda p(e^t - 1) + \frac{\lambda^2 p^2}{2n}(e^{2t} - 1) + \frac{\lambda^3 p^3}{3n^2}(e^{3t} - 1) + \dots$$

فاذن

$$\lim_{n \to \infty} K(t, n) = \lambda(e^t - 1) = K(t)$$

$$\therefore M(t) = e^{K(t)} = e^{\lambda(e^t - 1)}$$

لاحظ ان $M(t)$ تمثل الدالة المولدة لعزوم توزيع پواسون بالمعلمة λ . فاذن نستنتج ان التوزيع المقيد الى Y_n هو توزيع پواسون بالمعلمة λ .

تمارين الفصل الثامن

٨ ـ ١ : جد حجم العينة المطلوب سحبها من مجتمع توزيعه $N(\mu, 9)$ لغرض تقدير μ بحيث ان احتمال الفرق المطلق ، بين $\bar{X}, \mu$ لايزيد عن 0.8 ، لايقل عن 0.95 .

٨ ـ ٢ : افرض ان $X \sim N(10, 9)$. جد الحد الاعلى لاحتمال ان $|X - \mu| \geq \mu - \sigma$ ثم قارن هذا الاحتمال مع الاحتمال الحقيقي لحدوث الحادثة $\{|X - \mu| \geq \mu - \sigma\}$

(٨ ـ ٣ : افرض ان $X \sim b(1000, 0.5)$. جد الحد الاعلى لاحتمال ان $|X - np| \geq V(X)/10$. ثم قارن هذا الاحتمال مع الاحتمال الحقيقي لهذه الحادثة .

٨ ـ ٤ : ليكن X متغير عشوائي يسلك وفق دالة التوزيع المنتظم على الفترة $(-\sqrt{3}, \sqrt{3})$. جد الحد الاعلى لاحتمال ان $|X - \mu| \geq 3\sigma/2$ ثم قارن هذا الاحتمال بالاحتمال الحقيقي لحدوث هذه الحادثة .

٨ ـ ٥ : ليكن X متغير عشوائي بحيث ان $EX^2 = 13, EX = 3$. جد الحد الادنى الى $P_r(-2 < X < 8)$

٨ ـ ٦ : افرض ان $X \sim Gamma(2,4)$ وان $\bar{X}$ يمثل الوسط الحسابي لعينة عشوائية قوامها 128 مفرده مسحوبة من هذا التوزيع . جد قيمة تقريبية الى $P_r(7 < \bar{X} < 9)$.

٨ ـ ٧ : افرض ان $X \sim b(400, 0.2)$. جد قيمة تقريبية الى $P_r(X/n \geq 0.25)$.

٨ ـ ٨ : افرض ان $\bar{X}$ يمثل الوسط الحسابي لعينة عشوائية مسحوبة من توزيع كاما بالمعلمتين $\beta = 1, \alpha$. بين ان التوزيع المقيد الى الدرجة المعيارية Z_n المقابلة الى $\bar{X}$ هو $N(0,1)$.

٨ ـ ٩ : ليكن $X \sim Gamma(n, \beta)$ حيث ان β لاتعتمد على n ، وافرض ان $Y_n = X/n$. جد التوزيع المقيد الى Y_n..

٨ ـ ١٠ : افرض ان X_n يتوزع كتوزيع پواسون بالمعلمة n . وافرض ان $Y_n = (X_n - n)/\sqrt{n}$. بين ان الدالة المولدة لعزوم Y_n تتقارب من الدالة المولده لعزوم $N(0,1)$.

٨ ـ ١١ : اذا كان X يتوزع كتوزيع پواسون بالمعلمة $\lambda = 100$. جد قيمة تقريبية $P_r(85 \le X \le 110)$.

٨ ـ ١٢ : افرض ان S^2 يمثل التباين لقياسات عينه عشوائية قوامها n مسحوبة من $N(\mu, \sigma^2)$. برهن ان $nS^2 / n - 1$ يتقارب تصادفياً من σ^2 .





المعاينة من مجتمع طبيعي وتوزيعات المعاينة

الفصل التاسع

المعاينة من مجتمع طبيعي وتوزيعات المعاينة

Sampling from Normal Population & Sampling distributions

لاحظنا في فقرات الفصل الثامن مدى اهمية ودور التوزيع الطبيعي في تطبيقات النظرية الاحصائية . ان مبرهنة الغاية المركزية كافية لوحدها اثبات دور هذا التوزيع من خلال ملاحظتنا ان كافة التوزيعات الاحتمالية متقطعة كانت ام مستمرة تتقارب من التوزيع الطبيعي عندما يكون حجم العينة كبير . في هذا الفصل يبرز دور آخر مهم لهذا التوزيع وهو امكانية اشتقاق توزيعات احتمالية اخرى ذات اهمية تطبيقية كبيرة تسمى « توزيعات المعاينة » التي تفترض مسبقاً وجود معاينة من توزيع طبيعي . ان اهمية توزيعات المعاينة تبرز وبشكل واضح عند دراسة موضوعي اختبار الفرضيات الاحصائيةوبناء حدود الثقة (او فترات الثقة) لمعلمة (او مجموعة معالم) معينة . هذه التوزيعات تشترك بميزة واحدة هي ان معلمة (او معالم) دوالها الاحتمالية تسمى درجات الحرية degrees of freedom .

٩ ـ ١ : توزيع مربع كاي Chi- Square distribution

٩ ـ ١ ـ ١ : تعريف :

ليكن Y متغير عشوائي (او مؤشر احصائي) ذا توزيع $N(\mu, \sigma^2)$ ، وذلك يعني ان $Z=(Y-\mu)/\sigma \sim N(0,1)$ عند ئذ يقال ان Z^2 يمتلك توزيع مربع

كاي بدرجة حرية(*) واحدة. وبشكل اكثر عمومية وبفرض ان $Y_1, Y_2, \ldots, Y_n$ متغيرات عشوائية مستقلة بحيث ان $Y_i \sim N(\mu_i, \sigma_i^2)$ فان $Z_i = (Y_i - \mu_i)/\sigma_i \sim N(0, 1)$ وان Z_i^2 سيتوزع كتوزيع مربع كاي بدرجة حرية واحدة، وعندئذ يقال ان $\sum_{i=1}^{n} Z_i^2$ متغير عشوائي يتوزع كتوزيع مربع كاي بـ n درجة حرية. وكما سنلاحظ من الاشتقاق التالي فان توزيع مربع كاي يمثل حالة خاصة من عائلة توزيعات كاما عندما $\alpha = n/2$, $\beta = 2$.

٩ ـ ١ ـ ٢ : اشتقاق دالة توزيع مربع كاي

يمكن استنتاج دالة توزيع مربع كاي باستخدام اسلوب الدالة المولدة للعزوم وكمايلي : افرض ان $Z_1, Z_2, \ldots, Z_n$ متغيرات عشوائية مستقلة بحيث ان $Z_i \sim N(0, 1), \forall_i = 1, 2, \ldots, n$ وافرض ان $X = \sum_{i=1}^{n} Z_i^2$ لتكن $M_X(t)$ تمثل الدالة المولدة لعزوم X وان هذه الدالة موجودة لجميع قيم t المعرفة في الفترة $(-h, h)$ $h > 0$ وذلك يعني ان :

$$M_X(t) = Ee^{tX} = Ee^{t\sum_{i=1}^{n} Z_i^2} = E\prod_{i=1}^{n} e^{tZ_i^2}$$

(*) لابد لنا من اعطاء توضيح لمفهوم درجات الحرية كونه سوف يتكرر كثيراً في الفقرات اللاحقة. لتكن $X_1, X_2, \ldots, X_n$ قياسات عينة عشوائية قوامها n مسحوبة من مجتمع احصائي . وليكن $\bar{X}$ يمثل الوسط الحسابي لهذه القياسات. كما هو معلوم فان $\Sigma(X_i - \bar{X}) = 0$ كاحدى خصائص الوسط الحسابي فاذا كان معلوم لدينا $(n - 1)$ من هذه الانحرافات فان الانحراف الاخير سيتحدد تلقائياً، وذلك يعني اننا « احرار » في انتقاء $(n - 1)$ من هذه الانحرافات « ومقيدين » في انتقاء الاخير. فمثلاً اذا كان حجم العينة هو 5 متوسطها هو 4 ومعلوم لدينا اربعة قياسات فقط هي 2,3,4,5 فان القياسة الاخيرة يجب ان تكون العدد 6 لاغيره الذي يجعل مجموع الانحرافات عن الوسط مساوٍ للصفر. من هذا المثال البسيط فان عدد درجات الحرية المتاحة هو 4 = 1 ـ 5 . وعلى نحو اكثر ملائمة ان عملية فقدان مفردة واحدة ناجم عن قيامنا بتقدير متوسط هذا المجتمع μ من خلال حساب $\bar{X}$ اي ان $\bar{X}$ اصبح قيداً مفروضاً على هذه العينة. لذلك فان عدد درجات الحرية في هذه الحالة هو عدد مفردات العينة مطروحاً منه واحد. وبشكل عام اذا كان عدد مفردات العينة (او عدد المتغيرات العشوائية المستقلة) هو n وهنالك K من التقديرات المستقلة المطلوبة من هذه العينة (قيود على العينة) فان درجة الحرية المتاحة هي $d.f = n - k$ بشرط ان $K < n$.

وحيث ان Z_i متغيرات عشوائية مستقلة فان :

$$M_X(t) = \prod_{i=1}^{n} Ee^{tZ_i^2}$$

لكن $Z_i \sim N(0,1)$ فاذن :

$$Ee^{tZ_i^2} = \int_{-\infty}^{\infty} e^{tz_i^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z_i^2} dz_i$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} z_i^2 (1-2t)} dz_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \cdot \frac{Z_i^2}{(1-2t)^{-1}}} i \ldots (*)$$

لكن وبشكل عام لاحظنا في موضوع التوزيع الطبيعي لمتغير عشوائي مثل X ان :

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \sqrt{2\pi}\,\sigma \qquad \ldots(**)$$

وبتشبيه (*) مع (* *) نلاحظ ان :

$$\mu_z = 0 , \sigma_z^2 = (1-2t)^{-1}$$

وهذا يعني ان

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2} \cdot \frac{Z_i^2}{(1-2t)^{-1}}} = \sqrt{2\pi}\,(1-2t)^{-\frac{1}{2}}$$

$$\therefore Ee^{tZ_i^2} = (1-2t)^{-\frac{1}{2}}$$

وهذا يعني ان :

$$M_X(t) = \prod_{i=1}^{n} (1-2t)^{-\frac{1}{2}} = (1-2t)^{-\frac{n}{2}}$$

والصيغة الاخيرة تمثل الدالة المولدة لعزوم توزيع كاما بالمعلمتين $\alpha = \frac{n}{2}$ و $\beta = 2$ فاذن :

$$f\left(x ; \frac{n}{2}, 2\right) = \frac{1}{\Gamma\left(\frac{n}{2}\right) . 2^{n/2}} . x^{\frac{n}{2}-1} . e^{-\frac{x}{2}} \quad ; x > 0$$

ان الدالة الاخيرة تسمى دالة توزيع مربع كاي بـ n درجة حرية ، ويلاحظ انها حالة خاصةمن توزيع كاما . واستناداً لما تقدم يمكن صياغة تعريف متكامل لتوزيع مربع كاي وفق الاتي :

يقال ان المتغير العشوائي المستمر X يتوزع كتوزيع مربع كاي بـ n درجة حرية اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي :

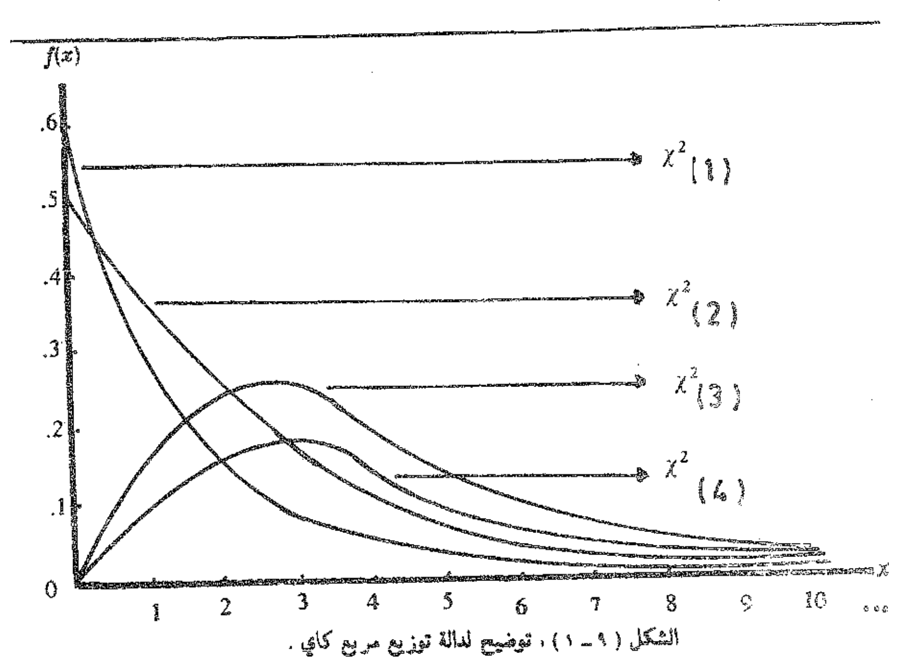
$$f(x ; n) = \frac{1}{\Gamma\left(\frac{n}{2}\right) . 2^{n/2}} . x^{\frac{n}{2}-1} . e^{-\frac{x}{2}} \quad ; x > 0$$

n : عدد موجب صحيح

$$= 0 \qquad \text{other wise}$$

ان الترميز الشائع لهذا التعريف هو $X \sim \chi^2_{(n)}$.

والشكل (٩ ـ ١) يوضح مخطط دالة هذا التوزيع عند درجات حرية مختلفة .



الشكل (٩ ـ ١) ، توضيح لدالة توزيع مربع كاي .

٩ ـ ١ ـ ٢ : الدالة التوزيعية لتوزيع χ^2

هنالك صيغ عديدة للدالة التوزيعية لتوزيع χ^2 تتباين فيما بينها من حيث الدقة في تحديد القيم الجدولية لهذا التوزيع . وفيما يلي عرض لصيغتين من هذه الصيغ :

أ ـ الدالة التوزيعية باستخدام توزيع پواسون .

بفرض ان $X \sim \chi^2_{(n)}$ عندئذٍ

$$F(x) = 1 - e^{-\frac{x}{2}} \cdot \sum_{j=0}^{r-1} \frac{(x/2)^j}{j!}, \; n = 2r$$ بحيث n عدد زوجي

$$= 1 - e^{-\frac{x}{2}} \cdot \sum_{j=0}^{r-2} \frac{(x/2)^{j+\frac{1}{2}}}{\Gamma\left(j+\frac{3}{2}\right)} - 2(1 - \Phi\sqrt{x}\,))$$

n عدد فردي بحيث $n=2r-1$ ، وان $\Phi(\sqrt{x}\,)$ تعني الدالة التوزيعية الى $N(0,1)$ عند $Z = \sqrt{x}$

مثال (١) اذا علمت ان $X \sim \chi^2_{(4)}$ جد $P_r(X \le 11{\cdot}1433)$.

الحل :

حيث أن n عدد زوجي وان $r = 2$ عليه فان

$$F(11{\cdot}1433) = 1 - e^{-\frac{11\cdot1433}{2}} \sum_{j=0}^{1} \frac{(11{\cdot}1433/2)^j}{j!}$$

$$= 1 - e^{-5\cdot57165} \cdot (1 + 5{\cdot}57165) = 0{\cdot}9750002 \simeq 0{\cdot}975$$

مثال (٢) : اذا علمت ان $X \sim \chi^2_{(5)}$ جد $P_r(X \leq 2{\cdot}6746)$.

الحل :

حيث n عدد فردي وان $r' = 3$، عليه فان

$$F(2{\cdot}6746) = 1 - e^{-\frac{2{\cdot}6746}{2}} \cdot \sum_{j=0}^{1} \frac{(2{\cdot}6746/2)^{j+\frac{1}{2}}}{\Gamma\left(j+\frac{3}{2}\right)} - 2(1 - \Phi(\sqrt{2{\cdot}6746}))$$

$$= 1 - e^{-1{\cdot}3373} \cdot (1{\cdot}3048765 + 1{\cdot}1633411) - 0{\cdot}101$$

حيث ان $\Phi(\sqrt{2{\cdot}6746}) = 0{\cdot}9495$ فاذن $F(2{\cdot}6746) = 0{\cdot}2509606 \simeq 0{\cdot}25$

ب ـ الدالة التوزيعية باستخدام تقريب Wilson - Hilferty

اقترح كل من Wilson - Hilferty عام ١٩٣١ التقريب التالي لحساب الدالة التوزيعية لتوزيع χ^2

$$F(x) \simeq \Phi\left\{\left[\left(\frac{x}{n}\right)^{\frac{1}{3}} - 1 + \frac{2}{9n}\right] \cdot \left(\frac{9n}{2}\right)^{\frac{1}{2}}\right\}$$

حيث n تمثل عدد درجات الحرية و x قيمة معطاة الى X وان $\Phi(.)$ تعني الدالة التوزيعية في $N(0,1)$. فمثلاً اذا كان $X \sim \chi^2_{(4)}$ فان

$$F(11{\cdot}1433) \simeq \Phi\left\{\left[\left(\frac{11{\cdot}1433}{4}\right)^{\frac{1}{3}} - 1 + \frac{2}{36}\right] \cdot \left(\frac{36}{2}\right)^{\frac{1}{2}}\right\}$$

$$= \Phi\{1{\cdot}9627847\} = 0{\cdot}975 .$$

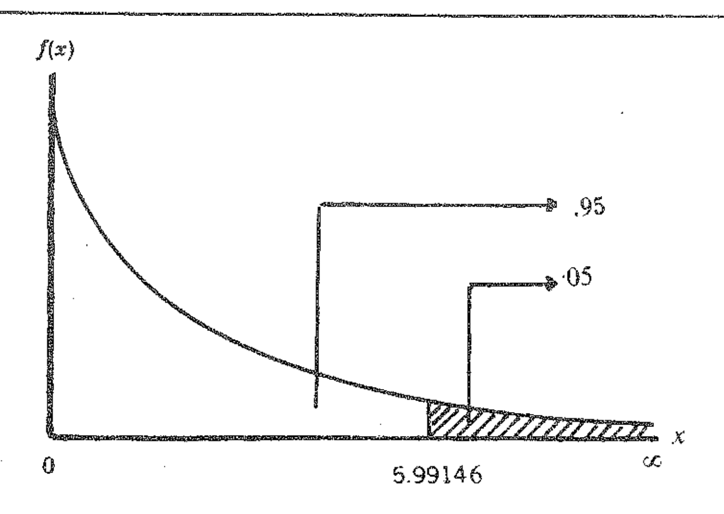
وعلى اساس الدالة التوزيعية (وفق اية صيغة كانت) تم تكوين جداول خاصة بهذا التوزيع تبين قيم χ^2 النظرية عند احتمال متراكم معين بالاستناد لعدد درجات الحرية n (لاحظ الجدول ٦ ملحق ب) . فمثلاً قيمة χ^2 النظرية عند درجة حرية 10 التي تعطي احتمالاً متراكماً مقداره 0.05 هي 3.9403 ، او انها قيمة χ^2 النظرية التي تجعل $P_r(\chi^2 \geq 3{\cdot}9403) = 0{\cdot}95$ كذلك فان الرمز الشائع لقيم χ^2 النظريـة هـو $\chi^2_n(\alpha)$ اي قيمـة χ^2 النظرية عند درجة حرية n ومستوى معنوية α اي انها تلك القيمة من قيم χ^2 المعرفة في الفترة $(0, \infty)$ التي تحقق :

$$P_r(\chi^2 \geq \chi^2_n(\alpha)) = \alpha \,,\, P_r(\chi^2 \leq \chi^2_n(\alpha)) = 1 - \alpha \,,\, 0 < \alpha < 1 \,.$$

فمثلا قيمة $\chi^2_2(0{\cdot}05)$ هي 5·99146 بحيث أن :

$$P_r(\chi^2 \geq 5{\cdot}99146) = 0{\cdot}05 \,,\, P_r(\chi^2 \leq 5{\cdot}99146) = 0{\cdot}95$$

والشكل (٩ ــ ٢) يوضح ذلك :



الشكل (٩ ــ ٢) ، توضيح لحساب $\chi^2_2(0{\cdot}05)$.

٩ ـ ١ ـ ٤ : الدالة المولدة لعزوم توزيع χ^2

لاحظنا في الفقرة (٩ ـ ٢ ـ ٢) ان توزيع χ^2 هو حالة خاصة من توزيع كاما عندما $\alpha = \frac{n}{2}, \beta = 2$. وهذا يعني انه يمكن وبسهولة استنتاج عزوم هذا التوزيع بمجرد التعويض عن α, β بما يساويهما في صيغ عزوم توزيع كاما وكما هو موضح بالجدول التالي :

الصيغة المطلوبة	توزيـــع كاما	توزيع χ^2
$M_X(t)$	$(1 - \beta t)^{-\alpha}$	$(1 - 2t)^{-n/2}$
الوسط EX	$\alpha . \beta$	n
التباين $V(X)$	$\alpha . \beta^2$	$2n$
EX^r	$\beta^r \dfrac{\Gamma(\alpha + r)}{\Gamma(\alpha)}$	$2^r . \dfrac{\Gamma\left(\frac{n}{2} + r\right)}{\Gamma\left(\frac{n}{2}\right)}$

ويلاحظ من الجدول السابق ان الوسط لتوزيع مربع كاي يتمثل بعدد درجات الحرية للتوزيع في حين ان تباين هذا التوزيع هو ضعف عدد درجات حريته . وهذا يعني ان متوسط هذا التوزيع هو دائماً اقل من تباينه .

٩ ـ ١ ـ ٥ : خاصية الجمع في توزيع χ^2.

افرض ان $X_1, X_2, \ldots, X_k$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim \chi^2_{(n_i)}$ عندئذ فان $Y = \sum_{i=1}^{k} X_i \sim \chi^2\left(\sum_{i=1}^{k} n_i\right)$.

البرهان : افرض ان الدالة المولدة لعزوم Y موجودة,فاذن

$$M_Y(t) = Ee^{tY} = Ee^{t\sum_{i=1}^{k} X_i}$$

$$= \prod_{i=1}^{k} M_{X_i}(t)$$

لكن

$$M_{X_i}(t) = (1-2t)^{-\frac{n_i}{2}}$$

فاذن

$$M_Y(t) = \prod_{i=1}^{k} (1-2t)^{-\frac{n_i}{2}} = (1-2t)^{-\frac{1}{2}\sum_{i=1}^{k} n_i}$$

ويلاحظ ان الصيغة الاخيرة تمثل الدالة المولدة لعزوم توزيع χ^2 بـ $\sum_{i=1}^{k} n_i$ درجة حرية فاذن $Y = \sum_{i=1}^{k} X_i \sim \chi^2_{\left(\sum_{i=1}^{k} n_i\right)}$.

فمثلاً اذا كان $X_3 \sim \chi^2_{(10)}, X_2 \sim \chi^2_{(6)}, X_1 \sim \chi^2_{(4)}$ وان هذه المتغيرات مستقلة تصادفياً عندئذٍ $Y = X_1 + X_2 + X_3 \sim \chi^2_{(20)}$.

٩ ـ ١ ـ ٦ : المنوال ونقاط الانقلاب في منحنى دالة توزيع χ^2

لاحظنا لدى دراستنا لتوزيع كاما ان المنوال كان قيمة x المساوية الى $\beta(\alpha - 1)$. الان بجعل $\beta = 2, \alpha = n/2$ نحصل على صيغة المنوال في توزيع χ^2 وهي $x = n-2$ وان

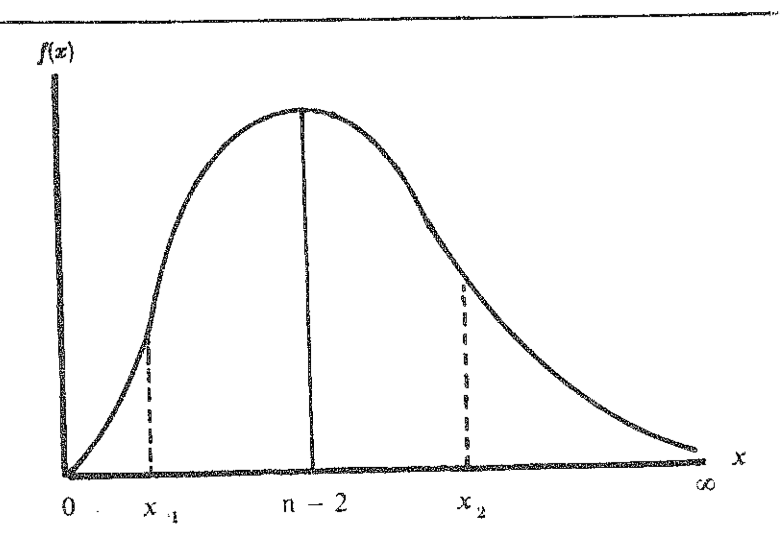
$$\text{Max.} f(x) = f(x)]_{x=n-2} = \frac{\left[e^{-1}\left(\frac{n}{2}-1\right)\right]^{\frac{n}{2}-1}}{2\Gamma\left(\frac{n}{2}\right)}$$

كذلك يمكن البيان ان المنحنى دالة هذا التوزيع نقطتي انقلاب تقعان على بعد متساوي الى يمين ويسار موقع المنوال . هاتين النقطتين هما :

$$x = (n-2) \pm \sqrt{2(n-2)}$$

$$= \text{Mode} \pm \sqrt{2. \text{Mode}}$$

والشكل (٩ ــ ٣) يوضح موقع المنوال ونقطتي الانقلاب .



(الشكل (٩ ــ ٣) ، توضيح لموقع المنوال ونقطتي الانقلاب في منحنى دالة توزيع مربع كاي .

٩ ــ ١ ــ ٧ : الالتواء لتوزيع χ^2 .

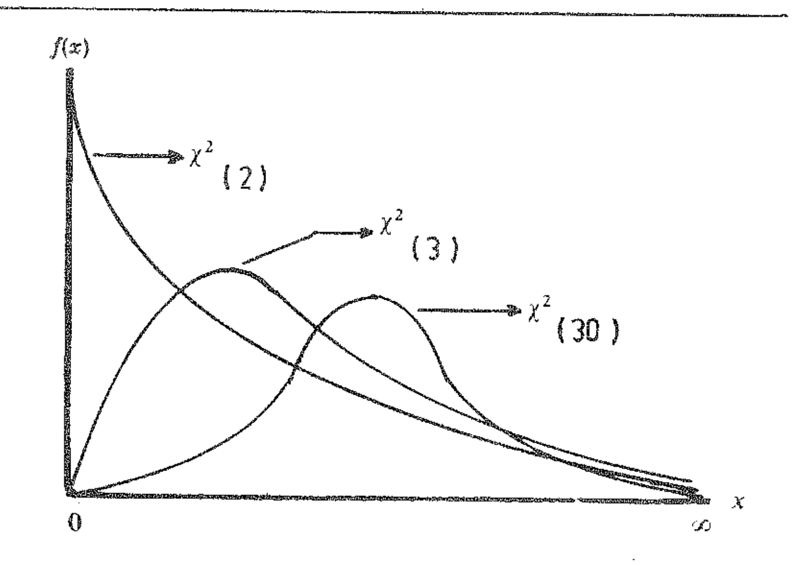
لاحظنا مما تقدم ان متوسط توزيع مربع كاي هو n وان تباينه هو 2n وان المنوال هو n − 2 . وهذا يعني ان معامل التواء هذا التوزيع وفق هذه المعطيات هو :

$$S_K = \frac{\text{mean} - \text{mode}}{\sqrt{\text{variance}}} = \frac{n - (n-2)}{\sqrt{2n}} = \sqrt{\frac{2}{n}} > 0$$

وهذا يعني ان منحنى توزيع مربع كاي هو دائماً ذات التواء موجب .
كذلك فان :

$$\lim_{n \to \infty} S_k = \lim_{n \to \infty} \sqrt{\frac{2}{n}} = 0$$

وهذا يعني انه بزيادة عدد درجات الحرية يبدأ منحنى دالة هذا التوزيع بالاقتراب من حالة التماثل . وسوف نبرهن في الفقرة اللاحقة ان توزيع مربع كاي يتقارب من التوزيع الطبيعي عندما $n \to \infty$. لاحظ الشكل (٩ ـ ٤) :



الشكل (٩ ـ ٤) ، توضيح لاقتراب منحنى توزيع مربع كاي من حالة التماثل

٩ ـ ١ ـ ٨ : خاصية التقارب لتوزيع χ^2 عندما تكون n عدد كبير

ليكن $X \sim \chi^2_{(n)}$ وان $Z = (X - n) / \sqrt{2n}$ تمثل الدرجة المعيارية المقابلة الى X . عندئذ فان $Z \sim N(0, 1)$ عندما $n \to \infty$.

البرهان : افرض ان $M_Z(t)$ موجودة فاذن

$$M_Z(t) = Ee^{tZ} = e^{-\frac{nt}{\sqrt{2n}}} . M_X\left(\frac{t}{\sqrt{2n}}\right)$$

لكن

$$M_X\left(\frac{t}{\sqrt{2n}}\right) = \left(1 - \frac{2t}{\sqrt{2n}}\right)^{-\frac{n}{2}}$$

فاذن

$$M_Z(t) = e^{-\sqrt{\frac{n}{2}}t}\left(1 - \sqrt{\frac{2}{n}}t\right)^{-\frac{n}{2}}$$

$$\therefore\ K_Z(t) = \ln M_Z(t) = -\sqrt{\frac{n}{2}}t - \frac{n}{2}\ln\left(1 - \sqrt{\frac{2}{n}}t\right)$$

$$;\quad \sqrt{\frac{2}{n}}t < 1$$

وحسب متسلسلة تايلر فان

$$\ln\left(1 - \sqrt{\frac{2}{n}}t\right) \quad -\sqrt{\frac{2}{n}}t - \frac{1}{2}\left(\sqrt{\frac{2}{n}}t\right)^2 -$$

$$\frac{1}{3}\left(\sqrt{\frac{2}{n}}t\right)^3 - \dots$$

$$= -\sqrt{\frac{2}{n}}\,t - \frac{t^2}{n} - \left(\sqrt{\frac{2}{n}}\right)^3 \cdot \frac{t^3}{3} - \ldots$$

$$\therefore\ K_Z(t) = -\sqrt{\frac{n}{2}}\,t - \frac{n}{2}\left(-\sqrt{\frac{2}{n}}\,t - \frac{t^2}{n} - \left(\sqrt{\frac{2}{n}}\right)^3 \cdot \frac{t^3}{3} \ldots\right)$$

$$= -\sqrt{\frac{n}{2}}\,t + \sqrt{\frac{n}{2}}\,t + \frac{t^2}{2} + \sqrt{\frac{2}{n}} \cdot \frac{t^3}{3}$$

$$+ 0\left(\frac{1}{n}\right)$$

حيث ان$0\left(\frac{1}{n}\right)$تعني حدود لاحقة تتضمن n في مقاماتها بقوى عليا .

$$\therefore\ K_Z(t) = \frac{t^2}{2} + \frac{t^3}{3} \cdot \frac{\sqrt{2}}{\sqrt{n}} + 0\left(\frac{1}{n}\right)$$

$$\therefore \lim_{n \to \infty} K_Z(t) = \frac{t^2}{2}$$

وهذا يعني ان

$$\lim_{n \to \infty} M_Z(t) = e^{\frac{1}{2}t^2}$$

والصيغة الاخيرة تمثل الدالة المولدة لعزوم $N(0,1)$ ، وهذا يعني ان $Z \sim N(0,1)$ عندما $n \to \infty$.اي ان توزيع مربع كاي يتقارب من التوزيع الطبيعي لدرجات حرية كبيرة .

ان لهذه الخاصية اهمية كبيرة من الناحية التطبيقية . فهي تسمح باستخدام جداول التوزيع الطبيعي كبديل لجداول توزيع مربع كاي (التي غالباً ماتكون منتهية بدرجة حرية قدرها 30) في حساب احتمال معين . وقد لوحظ عملياً انه عندما $n > 30$ فان توزيع مربع كاي يقترب من التوزيع الطبيعي وتزداد درجة

الاقتراب بزيادة عدد درجات الحرية . فمثلاً اذا كان $n > 30 , X \sim \chi^2_{(n)}$ عندئذٍ :

$$P_r (X \leq x) \simeq P_r \left(\frac{X - n}{\sqrt{2n}} \leq \frac{x - n}{\sqrt{2n}} \right) \simeq P_r (Z \leq z)$$

حيث $Z \sim N (0 , 1)$ ان هذا التقريب ليس دقيق جداً بالرغم من كونه يفي بالغرض المطلوب.لذا فقد تمكن العالم Fisher عام ١٩٤٤ بعمل تقريب اكثر دقة الموضحة صيغته بالآتي :

$$P_r (X \leq x) \simeq \Phi (\sqrt{2x} - \sqrt{2n - 1})$$

حيث ان $\Phi(.)$ تعني الدالة التوزيعية في $N (0 , 1)$ فمثلاً اذا كان $X \sim \chi^2_{(200)}$ وتطلب الامر حساب $P_r (X \leq 190)$ فان :

$$P_r (X \leq 190) \simeq \Phi (\sqrt{380} - \sqrt{399}) = \Phi (- 0{\cdot}48) = 0{\cdot}3156$$

٩ ــ ١ ــ ٩ : توزيع الدرجة المعيارية لمتوسط عينة مسحوبة من $N (\mu , \sigma^2)$

افرض ان $X_1 , X_2 , \ldots , X_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع طبيعي بوسط μ وتباين σ^2 وافرض ان $\bar{X}$ يمثل الوسط الحسابي لهذه العينة . وكما هو معلوم فان $\bar{X} \sim N (\mu , \sigma^2 / n)$ وان $Z = (\bar{X} - \mu) / (\sigma / \sqrt{n}) \sim N (0 , 1)$ وهذا يعني ان $Z^2 \sim \chi^2_{(1)}$. من الناحية التطبيقية فان Z^2 يعد مؤشر احصائي لاختبار فرضية تتعلق بالوسط μ (المجهول) عندما يكون تباين المجتمع معلوم او مقدر على اساس عينه عشوائية كبيرة الحجم .

٩ ــ ١ ــ ١٠ : توزيع تباين عينة مسحوبة من $N (\mu , \sigma^2)$.

افرض ان $X_1 , X_2 , \ldots , X_n$ تمثل قياسات عينة عشوائية من $N (\mu , \sigma^2)$ وافرض ان $\bar{X}$ يمثل متوسط هذه العينة وان $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 / n - 1$ يمثل تباينها .

عندئذ فان $(n - 1) S^2 / \sigma^2 \sim \chi^2_{(n-1)}$ بفرض ان $\bar{X}$ مستقل عن S^2

البرهان : سوف نبرهن اولاً النظرية التالية التي تبين ان $\bar{X}$ مستقل عن S^2 ثم نعود للمطلوب برهنته في هذه الفقرة .

مبرهنه : لتكن $X_1, X_2, \ldots, X_n$ متغيرات عشوائية مستقلة بحيث ان $X_i \sim N(\mu, \sigma^2)$ وان $Z_1, Z_2, \ldots, Z_n$ تمثل الدرجات المعيارية المقابلة لهذه المتغيرات بحيث ان $Z_i = (X_i - \mu)/\sigma$ وان $Z_i \sim N(0, 1)$ وافرض ان $R = \sum_{i=1}^{n} (Z_i - \bar{Z})^2$ عندئذ فان R مستقل عن $\bar{Z}$

البرهان : حيث ان X_i مستقل عن X_j فان Z_i مُستقل عن Z_j ايضاً .

الان تأمل المصفوفة التالية $A = \{a_{ij}\}$ ذات مرتبة $n \times n$ التي عناصرها موصوفة حسب مايلي :

$$a_{in} = \frac{1}{\sqrt{n}} \quad \forall_i = 1, 2, \ldots, n \qquad \ldots *$$

$$a_{ij} = \frac{1}{\sqrt{j(j+1)}} \quad i = 1, 2, \ldots, j \quad , j < n \qquad \ldots **$$

$$a_{j(j+1)} = \frac{-j}{\sqrt{j(j+1)}} \text{ ,} j < n \qquad \ldots ***$$

$$a_{ij} = 0 \qquad i > j+1, j < n \qquad \ldots ****$$

فمثلاً اذا كانت هذه المصفوفة من مرتبة 4 × 4، اي ان $n = 4$، فان مواقع هذه العناصر حسب الوصف اعلاه ستكون كالآتي وفق التأشيرات الموضحة ازاء كل حاله :

$$A = \begin{bmatrix} ** & ** & ** & * \\ *** & ** & ** & * \\ **** & *** & ** & * \\ **** & **** & *** & * \end{bmatrix}$$

اي ان :

$$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/2 \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/2 \\ 0 & -2/\sqrt{6} & 1/\sqrt{12} & 1/2 \\ 0 & 0 & -3/\sqrt{12} & 1/2 \end{bmatrix}$$

ويمكن وبسهولة ملاحظة ان المصفوفة A متعامدة orthogonal اي ان $AA^T = A^TA = I_n$ ، ونترك للقاريء بيان ذلك وفق المثال اعلاه، حيث ان I_n تعني المصفوفة الاحادية identity matrix ذات مرتبة n .

ليكن المتجه العشوائي $W = [W_1 W_2 \ldots W_n]$ معرف بالشكل $W = ZA$ ، حيث ان $Z = [Z_1 Z_2 \ldots Z_n]$ اي ان التحويل $W = ZA$ يعني ان :

$$W_n = \sqrt{n}\,\bar{Z}, W_j = \frac{1}{\sqrt{j(j+1)}}\left(\sum_{i=1}^{j} Z_i - jZ_{j+1}\right), j < n$$

ويطلب من القاريء متابعة خطوات البرهان وفق معطيات المثال اعلاه .

وحيث المتغيرات $Z_1, Z_2, \ldots, Z_n$ مستقلة وان $Z_i \sim N(0,1)$ وان A مصفوفة متعامدة فذلك يعني ان $W_1, W_2, \ldots, W_n$ هي ايضاً متغيرات عشوائية مستقلة وان كل منها ذات توزيع $N(0,1)$ كذلك وحيث ان

$$WW^T = \sum_{i=1}^{n} W_i^2 = (ZA)(ZA)^T = ZAA^TZ^T = ZIZ^T = ZZ^T$$

$$= \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} (Z_i - \bar{Z} + \bar{Z})^2 = \sum_{i=1}^{n} [(Z_i - \bar{Z}) + \bar{Z}]^2$$

$$= \sum_{i=1}^{n} (Z_i - \bar{Z})^2 + n\bar{Z}^2 + 2\bar{Z} \sum_{i=1}^{n} (Z_i - \bar{Z})$$

$$= \sum_{i=1}^{n} (Z_i - \bar{Z})^2 + n\bar{Z}^2 \qquad ; \sum_{i=1}^{n} (Z_i - \bar{Z}) = 0$$

فاذن

$$\sum_{i=1}^{n} W_i^2 = \sum_{i=1}^{n} (Z_i - \bar{Z})^2 + n\bar{Z}^2$$

$$= \sum_{i=1}^{n} (Z_i - \bar{Z})^2 + W_n^2$$

$$\therefore R = \sum_{i=1}^{n} (Z_i - \bar{Z})^2 = \sum_{i=1}^{n} W_i^2 - W_n^2 = \sum_{i=1}^{n-1} W_i^2$$

وحيث ان R تعتمد على $W_1, W_2, \ldots, W_{n-1}$ في حين ان $\bar{Z}$ تعتمد فقط على W_n فذلك يعني ان R مستقلة عن W_n اي ان $\sum_{i=1}^{n} (Z_i - \bar{Z})^2$ مستقل عن $\bar{Z}$ وهذا يعني ان متوسط العينة مستقل عن تباينها .

نعود الان الى برهنة المطلوب في الفقرة (٩ ــ ١ ــ ١٠) .

البرهان : تأمل مجموع المربعات التالي : $Q = \sum_{i=1}^{n} (X_i - \mu)^2$

باضافة وطرح $\bar{X}$ داخل القوس نحصل على :

$$\sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} [(X_i - \bar{X}) + (\bar{X} - \mu)]^2$$

$$= \sum_{i=1}^{n} (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2$$

وبقسمة الطرفين على σ^2 نحصل على

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2=\frac{(n-1)S^2}{\sigma^2}+\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

لكن وفق ماتقدم فان

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2=\sum_{i=1}^{n}Z_i^2\sim\chi^2_{(n)}\quad ,Z_i\sim N(0,1)$$

وحسب الفقرة (٩ ـ ١ ـ ٩) فان

$$\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2\sim\chi^2_{(1)}$$

عليه وحسب خاصية الجمع في توزيع مربع كاي فان $\frac{(n-1)S^2}{\sigma^2}$ يجب ان يكون ذا توزيع مربع كاي بـ $(n-1)$ درجة حرية . عليه، فان $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2_{(n-1)}$ وهذا يقابل توزيع R في النظرية السابقة حيث لاحظنا ان $R=\sum_{i=1}^{n-1}W_i^2$ قد توزعت كتوزيع χ^2 بـ $(n-1)$ درجة حرية بسبب ان $W_i\sim N(0,1)$

وبفرض ان $m=n-1$ عندئذٍ :

$$f(S^2;\sigma^2,m)=\frac{1}{\Gamma\left(\frac{m}{2}\right).2^{m/2}}.\left(\frac{mS^2}{\sigma^2}\right)^{\frac{m}{2}-1}.e^{-\frac{mS^2}{2\sigma^2}}$$

$$=\frac{\left(\frac{m}{2\sigma^2}\right)^{\frac{m}{2}-1}}{2\Gamma\left(\frac{m}{2}\right)}.(S^2)^{\frac{m}{2}-1}.e^{-\frac{mS^2}{\sigma^2}}\;;S^2>0$$

كذلك فان :

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{n-1}{\sigma^2} ES^2 = n-1$$

$$\therefore ES^2 = \sigma^2$$

وان

$$V\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{(n-1)^2}{\sigma^4} V(S^2) = 2(n-1)$$

$$\therefore V(S^2) = \frac{2\sigma^4}{n-1}$$

٩ ـ ١ ـ ١١ : استخدامات توزيع χ^2

لتوزيع مربع كاي استخدامات كثيرة في تطبيقات النظرية الاحصائية في الجوانب العملية نذكر بعضا منها دون اللجوء الى تفاصيلها :

أ ـ اختبار فيما اذا كان تباين مجتمع طبيعي مساو لقيمة معطاة مثل σ_0^2 .

ب ـ اختبار حسن المطابقة goodness of fit اي اختبار الفرضية القائلة بان قياسات عينة عشوائية قد سحبت من مجتمع ذي توزيع احتمالي معين .

جـ ـ اختبار الاستقلالية بين الصفات او مايسمى في بعض الاحيان اختبار جداول التوافق .

د ـ اختبار تجانس عدة تقديرات مستقلة لتباين المجتمع σ^2 .

هـ ـ اختبار تجانس عدة تقديرات مستقلة لمعامل الارتباط في المجتمع .

و ـ اختبار وجود ازدواج خطي - multicollinearity بين المتغيرات المستقلة في نموذج انحدار خطي متعدد .

ز ـ بناء حدود الثقة لتباين المجتمع σ^2 .

مثال (٣) : اذا كان $X \sim \chi^2_{(1)}$. برهن ان $Z = \sqrt{X} \sim N(0,1)$

البرهان : سوف نستخدم اسلوب التحويل في ايجاد توزيع Z . حيث ان :

$$Z = \sqrt{X} \qquad \therefore X = Z^2 \rightarrow \left| \frac{dx}{dz} \right| = 2Z \; ; 0 < z < \infty$$

كذلك فان

$$f(x) = \frac{1}{\Gamma\left(\frac{1}{2}\right) . 2^{1/2}} . x^{-\frac{1}{2}} . e^{-\frac{x}{2}} = \frac{1}{\sqrt{2\pi}} . x^{-\frac{1}{2}} . e^{-\frac{x}{2}}$$

عليه فان

$$g(z) = f(x)]_{x=z^2} . \left| \frac{dx}{dz} \right| = \frac{1}{\sqrt{2\pi}} (Z^2)^{-\frac{1}{2}} . e^{-\frac{z^2}{2}} . 2Z$$

$$\therefore g(z) = \frac{2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \; ; z > 0$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \; ; -\infty < z < \infty$$

عليه فان $z = \sqrt{X} \sim N(0, 1)$.

مثال (٤) : اذا كان $X \sim \chi^2_{(n)}$ وان $Y = \ln X$. اشتق صيغة للدالة المولدة لعزوم Y حول نقطة الاصل .

الحل :

افرض ان $M_Y(t)$ موجودة ، وذلك يعني ان :

$$M_Y(t) = Ee^{tY} = Ee^{t \cdot \ln X} = Ee^{\ln X^t} = EX^t$$

$$= \frac{1}{\Gamma\left(\frac{n}{2}\right) . 2^{n/2}} \int_0^\infty x^t . x^{\frac{n}{2} - 1} . e^{-x/2} \, dx$$

$$= \frac{1}{\Gamma\left(\frac{n}{2}\right) . 2^{n/2}} \int_0^\infty x^{\left(\frac{n}{2} + t\right) - 1} . e^{-\frac{x}{2}} \, dx$$

والتكامل الاخير يمثل تكاملا كاما بالمعلمتين $\alpha = \frac{n}{2} + t$, $\beta = 2$ وان قيمة هذا التكامل هي $\Gamma\left(\frac{n}{2} + t\right) . 2^{\frac{n}{2} + t}$. عليه فان :

$$M_Y(t) = \frac{1}{\Gamma\left(\frac{n}{2}\right) . 2^{n/2}} . \Gamma\left(\frac{n}{2} + t\right) . 2^{\frac{n}{2} + t}$$

$$= 2^t . \frac{\Gamma\left(\frac{n}{2} + t\right)}{\Gamma\left(\frac{n}{2}\right)}$$

تمارين عن توزيع مربع كاي

٩ ـ ١ : اذا علمت ان $X_1 \sim \chi^2_{(3)}$ مستقل عن $X_2 \sim \chi^2_{(5)}$. يطلب اجراء مايلي :

أ ـ جد التوزيع الاحتمالي الى $Y = X_1 + X_2$ ثم ارسم مخطط دالة هذا التوزيع .

ب ـ جد الوسط والتباين ومعامل التواء توزيع Y .

جـ ـ جد . $P_r(Y \leq 4)$, $P_r(Y > 3)$, $P_r(2 < Y < 6)$.

د ـ جد قيمة y_0 التي تحقق مايلي :

$P_r(Y \geq y_0) = 0.05$, $P_r(Y \leq y_0) = 0.99$.

٩ ـ ٢ : اذا علمت ان $X \sim \chi^2_{(6)}$. جد قيمة $F(8)$ باستخدام الاسلوبين الموضحين في الفقرة (٩ ـ ١ ـ ٣) .

٩ ـ ٣ : افرض ان $X_1, X_2, ..., X_k$ متغيرات عشوائية مستقلة بحيث ان كل منها يتوزع وفق توزيع منتظم مستمر على الفترة $(0, 1)$. وافرض ان $P = \prod_{i=1}^{k} X_i$. برهن ان $Y = -2\ln P$ يتوزع كتوزيع مربع كاي بـ $2K$ درجة حرية .

٩ ـ ٤ : اذا كان $X \sim \chi^2_{(n)}$. برهن ان $Y = \sqrt{2X} \sim N(\sqrt{2n}, 1)$ وان $Z = \sqrt{2X} - \sqrt{2n} \sim N(0, 1)$ عندما $n \to \infty$.

٩ ـ ٥ : افرض ان $X_1, X_2, ..., X_n$ تمثل عينة عشوائية من $N(\mu_x, \sigma_x^2)$ وان $Y_1, Y_2, ..., Y_m$ تمثل عينة عشوائية من $N(\mu_y, \sigma_y^2)$. وبفرض ان المجتمعين مستقلين وان $\bar{X}, S_x^2$ هما الوسط والتباين للعينة الاولى وان $\bar{Y}, S_y^2$ هما الوسط والتباين للعينة الثانية ، وتحت الفرض القائل ان متوسط العينة مستقل عن تباينها يطلب اجراء مايلي :

أ ـ اشتق التوزيع الاحتمالي للمتغير $Z = \frac{(n-1)S_x^2}{\sigma_x^2} + \frac{(m-1)S_y^2}{\sigma_y^2}$

ب ـ اشتق الحالة العامة للتوزيع الاحتمالي الى Z في حالة وجود K من المجتمعات الطبيعية المستقلة .

٩ ـ ٦ : اذا كان $X_1 \sim \chi^2_{(n)}$ وان $X_2 \sim \chi^2_{(m)}$. برهن ان $Y = X_1 / (X_1 + X_2)$ يتوزع كتوزيع بيتابالمعلمتين $\alpha = \frac{1}{2} n , \beta = \frac{1}{2} m$.

٩ ـ ٧ : اذا علمت ان $X_1, X_2, \ldots, X_n$ تمثل عينة عشوائية من $N(\mu, 1)$. جد التوزيع الاحتمالي الى $Y = \sum_{i=1}^{n} (X_i - \mu)^2$. ثم جد الوسط والتباين للمتغير Y .

٩ ـ ٢ :توزيع ستودينت ـ ت ـ Student -t- distribution

يقترن مفهوم هذا التوزيع بالعينات الصغيرة (تطبيقاً $n < 30$) . ويعد العالم الاحصائي W. S. Gosset اول من كتب عن هذا التوزيع عام ١٩٠٨ تحت اسم مستعار هو Student ، في حين قام العالم الاحصائي R. A. Fisher (١٩٦٢ ـ ١٨٩٠) بعمل اضافات نوعية لهذا التوزيع المهم عام ١٩٢٠ .

وكما لاحظنا في الفقرة (٨ ـ ٢ ـ ٣) فان توزيع المتغيرات الاحصائية مثل $\sqrt{n}(X-\mu)/\sigma$, $(X-m)/\sqrt{m}$ وغيرها توزعت كتوزيع $N(0,1)$ عندما يكون حجم العينة كبير الامر الذي سمح لنا بامكانية استخدام جداول التوزيع الطبيعي ، لكن ماتقدم غير ممكن في حالة العينات الصغيرة بسبب ابتعاد توزيع هذه المتغيرات عن التوزيع الطبيعي الامر الذي يتطلب ايجاد توزيعاتها في مثل هذه الاحوال . وفيما يلي عرض لطريقة فيشر في اشتقاق دالة توزيع t .

٩ ـ ٢ ـ ١ : اشتقاق دالة توزيع t .

افرض ان $X_1 \sim N(0,1)$ مستقل عن $X_2 \sim \chi^2_{(n)}$. عندئذٍ فان النسبة $t = X_1/\sqrt{X_2/n}$ سوف تتوزع كتوزيع t بدرجة حرية n وفق دالة كثافة احتمالية

$$g(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \cdot \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < \infty$$

البرهان :

ان التوزيع المشترك للمتغيرين X_1, X_2 هو

$$f(x_1, x_2) = f(x_1) \cdot f(x_2)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2} \cdot \frac{1}{\Gamma\left(\frac{n}{2}\right) \cdot 2^{n/2}} \cdot x_2^{\frac{n}{2}-1} \; e^{-\frac{x_2}{2}}$$

$$= \frac{1}{\sqrt{2\pi} \cdot \Gamma\left(\frac{n}{2}\right) \cdot 2^{n/2}} \; x_2^{\frac{n}{2}-1} \cdot e^{-\frac{1}{2}(x_1^2 + x_2)}$$

افرض ان $Z = X_2$ ، فاذن :

$$t = \frac{X_1}{\sqrt{X_2/n}} = \frac{X_1}{\sqrt{Z/n}} \rightarrow X_1 = t\sqrt{\frac{Z}{n}}, X_2 = Z$$

وحيث ان مقام النسبة t ، اي $\sqrt{\frac{X_2}{n}}$ كمية موجبة وان $-\infty < x_1 \infty$ $-$ ∞

فذلك يعني ان $-\infty < t < \infty$ وان $z > 0$ طالما ان $x_2 > 0$. ان معامل التحويل (جاكوبيان) J هو :

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial t} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial t} & \frac{\partial x_2}{\partial z} \end{vmatrix} = \begin{vmatrix} \sqrt{\frac{z}{n}} & \frac{t}{2\sqrt{nz}} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{Z}{n}} \rightarrow |J| = \sqrt{\frac{z}{n}}$$

فاذن :

$$g(t, z) = f(x_1, x_2)\Big]_{\substack{x_1 = t\sqrt{\frac{z}{n}} \\ x_2 = z}} \cdot |J|$$

$$\therefore g(t,z) = \frac{1}{\sqrt{2\pi}\,.\,\Gamma\left(\frac{n}{2}\right).\,2^{n/2}}\,.\,z^{\frac{n}{2}-1}\,.\,e^{-\frac{1}{2}\left(\frac{t^2 z}{n}+z\right)}.\,\sqrt{\frac{Z}{n}}$$

$$= \frac{1}{\sqrt{n\pi}\,.\,\Gamma\left(\frac{n}{2}\right)}\left(\frac{1}{2}\right)^{\frac{n+1}{2}}.\,z^{\frac{n}{2}-\frac{1}{2}}\,.\,e^{-\frac{z}{2}\left(\frac{t^2}{n}+1\right)}$$

الان باجراء التكامل نسبة للمتغير Z يمكن الحصول على الدالة الاحتمالية الحدية للمتغير t . اي ان :

$$g(t) = \frac{\left(\frac{1}{2}\right)^{\frac{n+1}{2}}}{\sqrt{n\pi}\,.\,\Gamma\left(\frac{n}{2}\right)}\int_0^\infty z^{\left(\frac{n+1}{2}\right)-1}\,e^{-z.\frac{1}{2}\left(\frac{t^2}{n}+1\right)}\,dz$$

ويتضح ان التكامل اعلاه يمثل تكامل كاما بالمعلمتين

$$\beta = \frac{2}{\frac{t^2}{n}+1},\ \alpha = \frac{n+1}{2}$$

وبذلك فان ناتج هذا التكامل هو :

$$.\,\Gamma\left(\frac{n+1}{2}\right)\left(\frac{2}{\frac{t^2}{n}+1}\right)^{\frac{n+1}{2}}$$

عليه فان :

$$g(t) = \frac{1}{\sqrt{n\pi}\,.\,\Gamma\left(\frac{n}{2}\right)} \cdot \left(\frac{1}{2}\right)^{\frac{n+1}{2}} . \Gamma\left(\frac{n+1}{2}\right) .$$

$$\frac{2^{\frac{n+1}{2}}}{\left(\frac{t^2}{n}+1\right)^{\frac{n+1}{2}}}$$

$$\therefore g(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} . \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \; ; -\infty < t < \infty$$

وفي هذه الحالة يقال ان $g(t)$ تمثل دالة الكثافة الاحتمالية للمتغير العشوائي t بدرجة حرية n . وبالرموز يقال ان $t \sim t_{(n)}$.

اذا كانت $n = 1$ فان $g(t) = \pi^{-1}(1 + t^2)^{-1}$ والدالة $g(t)$ في هذه الحالة تمثل دالة توزيع كوشي بالمعلمتين $b = 1 , a = 0$. وهذا يعني ان توزيع كوشي هو حالة خاصة من توزيع t عندما $n = 1$.

٩ ـ ٢ ـ ٢ : الدالة التوزيعية لتوزيع .

لقد تم اقتراح العديد من الصيغ التقريبية الخاصة بالدالة التوزيعية لتوزيع t تباينت فيما بينها من حيث الدقة في حساب قيم t النظرية (اي تكوين الجداول الخاصة بهذا التوزيع) . ونستعرض ادناه صيغتين فقط منها وهي

أ ـ صيغة Fisher

اقترح العالم R.A.Fisher عام ١٩٢٥ الصيغة التالية لحساب قيم تقريبية لهذا التوزيع على اساس حساب الاحتمال المتراكم لغاية قيمة معطاة الى t مثل t_0 ، وهي :

$$F(t_0) = P_r(t_{(n)} \le t_0) = \int_{-\infty}^{t_0} g(t)\,dt$$

$$\simeq \Phi(t_0) - Z(t_0).(A + B + C)$$

حيث ان :

$\Phi(t_0)$ تعني الدالة التوزيعية لتوزيع $N(0,1)$ عند القيمة t_0 وان $Z(t_0)$ تعني قيمة دالة توزيع $N(0,1)$ عند القيمة t_0 .

$$A = \frac{t_0}{4n}(t_0^2 + 1),\ B = \frac{t_0}{96n^2}(3t_0^6 - 7t_0^4 - 5t_0^2 - 3),$$

$$C = \frac{t_0}{384n^3}(t_0^{10} - 11t_0^8 + 14t_0^6 + 6t_0^4 - 3t_0^2 - 15)$$

وقد بين فيشر ان مقدار الخطأ المطلق عند حساب $F(t)$ لايتجاوز 0.000005 . فمثلاً لتوزيع t بتسع درجات حرية فان قيمة $F(2{\cdot}821)$ يمكن حسابها وفق صيغة فيشر كالآتي :

ان $t_0 = 2{\cdot}821$ فاذن :

$$Z(2{\cdot}821) = \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}(2{\cdot}821)^2} = 0{\cdot}00746177,\ \Phi(2{\cdot}821) = 0{\cdot}9976$$

وان

$$A = 0{\cdot}7019620,\ B = 0{\cdot}3721626,\ C = -0{\cdot}0484073$$

فاذن

$$F(2{\cdot}821) = P_r(t_{(9)} \le 2{\cdot}821) = 0{\cdot}9899465 \simeq 0{\cdot}99$$

ب ــ صيغة Elfving

اقترح G. Elfving عام ١٩٥٥ الصيغة التالية لحساب قيم تقريبية لهذا التوزيع على اساس حساب الاحتمال المتراكم لغاية قيمة معطاة الى t مثل t_0 وهي :

$$F(t_0) = P_r(t_{(n)} \le t_0)$$

$$\simeq \Phi(\lambda t_0) + \frac{5}{96}\left[\frac{\lambda t_0^5}{n^2}\left(1 + \frac{t_0^2}{2n}\right)^{-\frac{n+4}{2}} - Z\left(\frac{\lambda t_0}{\sqrt{2}}\right)\right]$$

حيث ان $Z(g)$ تعني قيمة دالة توزيع $N(0, 1)$ عند القيمة g وان

$$\lambda = (n - 0.5)^{\frac{1}{2}} \cdot \left(n + \frac{1}{2} t_0^2\right)^{-\frac{1}{2}}$$

وقد بين Elfving أن مقدار الخطا الحاصل في حساب $F(t_0)$ لا يتجاوز المقدار $F(t_0)/2n^2$. فمثلاً لتوزيع t بتسع درجات حرية فان قيمة $F(2.821)$ يتم حسابها كالآتي :

$$\lambda = 0.8092607, Z\left(\frac{\lambda t_0}{\sqrt{2}} = 1.6142713\right) = 0.1084056,$$

$$\Phi(\lambda t_0 = 2.283) = 0.9887 \quad \therefore F(2.821) = 0.9916601 \simeq 0.99$$

لاحظ من هذا المثال ان مقدار الخطأ هو $\frac{1}{2n^2} \cdot F(2.821) = 0.00612135$.

وأيا كانت صيغة الدالة التوزيعية فقد تم بناء جداول خاصة بهذا التوزيع (لاحظ الجدول ٧ ملحق ب) تبين قيم t النظرية عند درجة حرية n التي تعطي احتمالاً متراكماً مقداره $F(t)$. فمثلاً قيمة t النظرية عند درجة حرية 6 والتي تعطي احتمالاً متراكماً مقداره 0.95 . هي 1.943 ، انظر الشكل (٩ ــ ٦) .

٩ ـ ٢ ـ ٣ : الوسط والتباين لتوزيع t .

من تعريف توزيع t لاحظنا ان :

$$t = \frac{X_1}{\sqrt{\frac{X_2}{n}}}, X_1 \sim N(0,1), X_2 \sim \chi^2_{(n)}$$

وان X_1 مستقل تصادفياً عن X_2 . فاذن

$$Et = E\left(\frac{X_1}{\sqrt{X_2/n}}\right) = EX_1 . E \frac{1}{\sqrt{X_2/n}}$$

لكن $EX_1 = 0$ ، عليه فان $Et = 0$. وهذا يعني ان متوسط توزيع t هو صفر وهو نفس متوسط $X_1 \sim N(0,1)$. كذلك فان

$$V(t) = Et^2 - (Et)^2 = Et^2$$

فاذن

$$Et^2 = E\left(\frac{X_1}{\sqrt{X_2/n}}\right)^2 = nEX_1^2 . E\frac{1}{X_2}$$

لكن $EX_1^2 = 1$ طالما ان $X_1 \sim N(0,1)$. وان

$$E\frac{1}{X_2} = \frac{1}{\Gamma\left(\frac{n}{2}\right).2^{n/2}}\int_0^\infty \frac{1}{x_2} x_2^{\frac{n}{2}-1} . e^{-x_2/2} dx_2$$

$$= \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}\int_0^\infty x_2^{\left(\frac{n}{2}-1\right)-1} . e^{-x_2/2}\, dx_2$$

$$= \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}} \cdot \Gamma\left(\frac{n}{2}-1\right) . 2^{\frac{n}{2}-1} = \frac{1}{\left(\frac{n}{2}-1\right)\cdot 2} = \frac{1}{n-2}$$

$$\therefore \; V(t) = \frac{n}{n-2} \; ; n > 2$$

ويلاحظ ان :

$$\lim_{n \to \infty} V(t) = \lim_{n \to \infty} \frac{n}{n-2} = \frac{1}{1-\frac{2}{n}} = 1 = V(X_1)$$

وسوف نلاحظ في فقرة لاحقة ان توزيع t يقترب من $N(0,1)$ عندما $n \to \infty$.

٩ ــ ٢ ــ ٤ : المنوال ونقاط الانقلاب في منحنى دالة توزيع t .

كما هو معلوم فان المنوال يمثل قيمة t الناتجة من حل المعادلة التفاضلية $g'(t)$ بشرط ان $g''(t) < 0$. وحيث ان

$$g(t) = k\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \quad , k = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{2\pi} \; . \Gamma\left(\frac{n}{2}\right)}$$

$$\therefore \log g(t) = \log k - \frac{n+1}{2} \log\left(1+\frac{t^2}{n}\right)$$

وباشتقاق الطرفين نسبة الى t نحصل على :

$$g'(t) = -g(t)\left[\frac{(n+1)t}{n+t^2}\right] \quad , g(t) > 0$$

وبجعل $g'(t)=0$ نحصل على $t=0$. وبايجاد المشتقة الثانية نحصل على ..

$$g''(t)=(n+1)g(t)\left[\frac{n-t^2}{(n+t^2)^2}-\frac{(n+1)t^2}{(n+t^2)^2}\right]$$

$$\therefore g''(t)\Big]_{t=0}=-\frac{(n+1)}{n}.k<0,k=g(0)>0$$

وحيث ان المشتقة الثانية للدالة $g(t)$ سالبة عندما $t=0$ فذلك يعني ان المنوال في توزيع t هو قيمة $t=0$..

ولغرض ايجاد نقاط انقلاب منحنى دالة توزيع t ، نجعل $g''(t)=0$ ومنها نحصل على

$$n-t^2-(n+1)t^2=0\rightarrow t=\pm\sqrt{\frac{n}{n+2}}=\pm t^*,t^*$$

وهذا يعني ان المنحنى دالة هذا التوزيع نقطتي انقلاب تقعان على بُعد متساوٍ الى يمين ويسار المنوال . كذلك فان

$$\lim_{n\to\infty} t^*=\pm\lim_{n\to\infty}\sqrt{\frac{n}{n+2}}=\pm\lim_{n\to\infty}\sqrt{\frac{1}{1+2/n}}=\pm 1$$

وهما نفس نقطتي انقلاب منحنى دالة $N(0,1)$.

اضافة لما تقدم فان منحنى دالة هذا التوزيع متماثل عند النقطة $t=0$. فاذا كانت $t_0>0$ قيمة من قيم t ، فان $g(-t_0)=g(t_0)$. وان

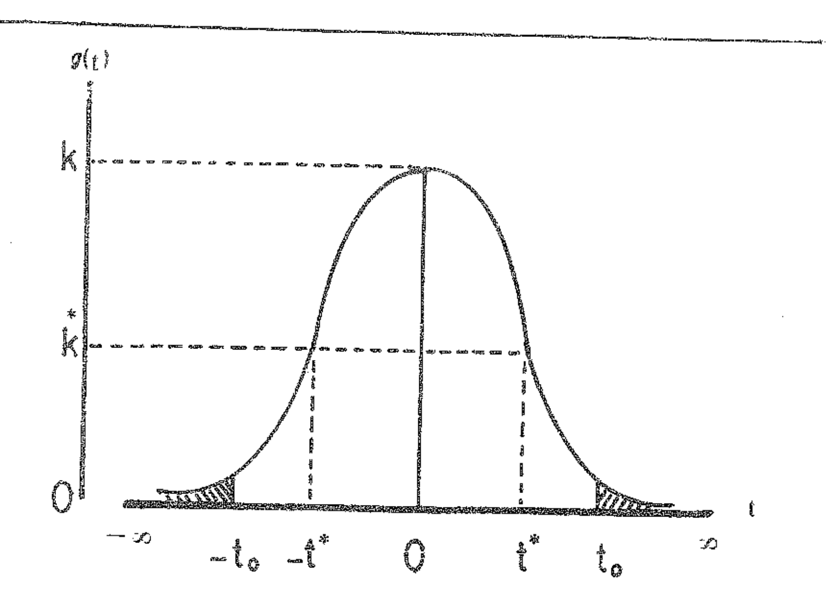
$$\text{Max}.\,g(t)=g(t)\Big]_{t=0}=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,.\,\Gamma\left(\frac{n}{2}\right)}=k$$

كذلك فان

$$g\left(t = -\sqrt{\frac{n}{n+2}}\right) = g\left(t = \sqrt{\frac{n}{n+2}}\right)$$

$$= k\left(\frac{n+3}{n+2}\right)^{-\frac{n+1}{2}} = K^*$$

والشكل (٩ ـ ٥) يوضح مخطط دالة هذا التوزيع .



الشكل (٩ ـ ٥) : توضيح لمخطط دالة توزيع .

ونلاحظ من الشكل (٩ ـ ٥) مايلي :

أ ـ ان $$P_r(t \le 0) = P_r(t \ge 0) = \frac{1}{2}$$

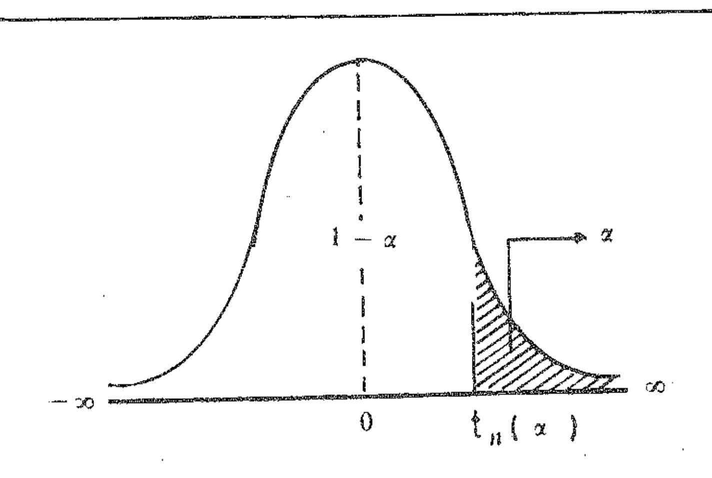
ب ـ ان $$P_r(t \le -t_0) = P_r(t \ge t_0)$$

جـ ـ ان $$P_r(t \le -t_0) = 1 - P_r(t \le t_0)$$

ووفق ماتم توضيحه في الفقرة (٩ ـ ٢ ـ ٢) فان قيم t النظرية غالباً مايرمز لها بالشكل $t_n(\alpha)$ اي قيمة t بدرجة حرية n ومستوى معنوية α . وهذا يعني ان $t_n(\alpha)$ تمثل قيمة t النظرية التي تحقق مايلي :

$$P_r(t_n \ge t_n(\alpha)) = \alpha \, , P_r(t_n \le t_n(\alpha)) = 1 - \alpha \, , 0 < \alpha < 1$$

وكما هو موضح في الشكل (٩ ـ ٦) :



الشكل (٩ ـ ٦) : توضيح لقيم t النظرية

٩ ـ ٢ ـ ٥ : عزوم توزيع t .

من الصعوبة جداً صياغة الدالة المولدة لعزوم توزيع t بشكل مألوف كما لاحظنا في التوزيعات السابقة . وحيث ان الهدف من هذه الدالة هو حساب عزوم التوزيع فاننا سوف نستعيض عنها من خلال ايجاد صيغة للعزوم ذي المرتبة r حول نقطة الاصل. وحيث ان هذا التوزيع متماثل فذلك يعني ان عزوم التوزيع حول

نقطة الاصل ذات مراتب فردية مساوية للصفر. اي ان $Et^{2r+1} = 0 \ \forall r = 0, 1, 2, \ldots$
كذلك وحيث أن $Et = 0$ فذلك يعني أن عزوم هذا التوزيع المركزية هي ذاتها عزومه حول نقطة الاصل. عليه فان العزم المركزي ذو مرتبة $2r$، $r = 1, 2, \ldots$ هو

$$Et^{2r} = \int_{-\infty}^{\infty} t^{2r} . g(t)\, dt = k \int_{-\infty}^{\infty} t^{2r} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt$$

$$= 2k \int_{0}^{\infty} t^{2r} . \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt$$

الان بفرض ان $1 + \frac{t^2}{n} = \frac{1}{Y}$ فاذن $t^2 = n(1-Y)/Y$. وعندما $t = 0$ فان $Y = 1$
وعندما $t \to \infty$ فان $Y \to 0$. فاذن $0 < y < 1$ كذلك فان

$$2t\, dt = -\frac{n}{Y^2}\, dY \quad \to dt = -\frac{n}{2t\, Y^2}\, dY$$

$$\therefore Et^{2r} = 2k \int_{1}^{0} t^{2r} . \left(1 + \frac{n(1-y)}{ny}\right)^{-\frac{n+1}{2}} \left(-\frac{n}{2t\, y^2}\, dy\right)$$

$$= nk \int_{0}^{1} (t^2)^{\frac{2r-1}{2}} . y^{\frac{n+1}{2}} . \frac{1}{y^2}\, dy$$

$$= nk \int_{0}^{1} (t^2)^{\frac{2r-1}{2}} . y^{\frac{n+1}{2} - 2}\, dy$$

$$= nk \int_0^1 \left(\frac{n(1-y)}{y} \right)^{\frac{2r-1}{2}} . y^{\frac{n+1}{2}-2} dy$$

$$= n^{r+\frac{1}{2}} . k \int_0^1 (1-y)^{r-\frac{1}{2}} . y^{-r+\frac{1}{2}} . y^{\frac{n+1}{2}-2} dy$$

$$= n^{r+\frac{1}{2}} . k \int_0^1 y^{\left(\frac{n}{2}-r\right)-1} . (1-y)^{\left(r+\frac{1}{2}\right)-1} dy$$

ويلاحظ ان التكامل الاخير يمثل تكامل بيتا بالمعلمتين $\alpha = \frac{n}{2} - r$ ، $\beta = r + \frac{1}{2}$. وان قيمة هذا التكامل هي :

$$\Gamma\left(\frac{n}{2}-r\right) . \Gamma\left(r+\frac{1}{2}\right) \left[\Gamma\left(\frac{n+1}{2}\right)\right]^{-1}$$

فاذن :

$$Et^{2r} = n^{r+\frac{1}{2}} . \frac{\Gamma\left(\frac{n+1}{2}\right) . \Gamma\left(\frac{n}{2}-r\right) . \Gamma\left(r+\frac{1}{2}\right)}{\sqrt{n\pi} . \Gamma\left(\frac{n}{2}\right) . \Gamma\left(\frac{n+1}{2}\right)}$$

$$= n^r . \frac{\Gamma\left(\frac{n}{2}-r\right) . \Gamma\left(r+\frac{1}{2}\right)}{\sqrt{\pi} . \Gamma\left(\frac{n}{2}\right)} , n > 2r, r = 1, 2, \ldots$$

وبشكل خاص اذا كانت $r = 1$ فاننا نكون بصدد حساب العزم المركزي الثاني اي تباين التوزيع المنوه عنه في الفقرة (٩ ـ ٢ ـ ٣) . كذلك عندما $r = 2$ فاننا نكون بصدد حساب العزم المركزي الرابع وهذا قيمته مساوية الى :

$$Et^4 = \frac{3n^2}{(n-2)(n-4)} \; ; n > 4$$

عليه فان معامل التفلطح في هذا التوزيع هو $\lambda_2 = \beta_2 - 3$ وان

$$\beta_2 = \frac{Et^4}{(Et^2)^2} = \frac{3(n-2)}{n-4}$$

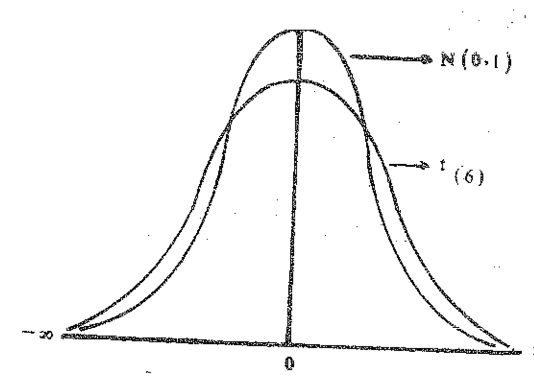
$$\therefore \lambda_2 = \frac{3(n-2)}{n-4} - 3 > 0, \quad \frac{n-2}{n-4} > 1$$

وهذا يعني ان منحنى دالة توزيع t اكثر تفلطحاً من منحنى دالة $N(0,1)$ ويتطابق المنحنيان في حالة n كبيرة اي ان :

$$\lim_{n \to \infty} \lambda_2 = \lim_{n \to \infty} \beta_2 - 3$$

$$= 3 \lim_{n \to \infty} \frac{n-2}{n-4} - 3 = 3 \lim_{n \to \infty} \frac{1 - \frac{2}{n}}{1 - \frac{4}{n}} - 3 = 0$$

والشكل (٩ ـ ٧) يوضح مقارنة بين توزيع t بست درجات حرية و $N(0,1)$.



الشكل (٩ ـ ٧) ، مقارنة بين توزيع $t_{(6)}$ وتوزيع $N(0,1)$.

٩ ـ ٢ ـ ٦ : الشكل التقاربي لتوزيع t .

ان توزيع t يتقارب من التوزيع الطبيعي المعياري عندما $n \rightarrow \infty$. اي ان

$$\lim_{n \to \infty} g(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}, -\infty < t < \infty$$

البرهان :

$$\lim_{n \to \infty} g(t) = \lim_{n \to \infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma\left(\frac{n}{2}\right)} \cdot \lim_{n \to \infty} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

لكن وبشكل عام ولاي عددين موجبين مثل m, k فان

$$\lim_{m \to \infty} \frac{\Gamma(m+k)}{\Gamma(m)} = \lim_{m \to \infty} \frac{(m+k-1)!}{(m-1)!} = \lim_{L \to \infty} \frac{(L+k)!}{L!},$$

$$L = m - 1$$

$$= \lim_{L \to \infty} \frac{(L+k)(L+k-1)\dots(L+1)L!}{L!}$$

$$= \lim_{L \to \infty} (L+k)(L+k-1) \quad \dots (L+1)$$

$$= \lim_{L \to \infty} (L+k)(L+k-1)\dots(L+1) \cdot \frac{L^k}{L^k}$$

$$= \lim_{L \to \infty} \left(1 + \frac{k}{L}\right)\left(1 + \frac{k-1}{L}\right)\dots\left(1 + \frac{1}{L}\right) . L^k = L^k$$

$$= L^k, L \rightarrow \infty$$

او ان

$$\lim_{m \to \infty} \frac{\Gamma(m+k)}{\Gamma(m)} = m^k , m \to \infty$$

الان لو اخترنا $m = \frac{n}{2}$, $k = \frac{1}{2}$ فذلك يعني ان

$$\lim_{n \to \infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = \left(\frac{n}{2}\right)^{\frac{1}{2}} ; n \to \infty$$

فاذن

$$\lim_{n \to \infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma\left(\frac{n}{2}\right)} = \frac{1}{\sqrt{n\pi}} \cdot \frac{\sqrt{n}}{\sqrt{2}} = \frac{1}{\sqrt{2\pi}}$$

وان

$$\lim_{n \to \infty} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} = \lim_{n \to \infty} \left[\left(1 + \frac{t^2}{n}\right)^n\right]^{-\frac{1}{2}}$$

$$\lim_{n \to \infty} \left(1 + \frac{t^2}{n}\right)^{-\frac{1}{2}}$$

$$= \left[\lim_{n \to \infty} \left(1 + \frac{t^2}{n}\right)^n\right]^{-\frac{1}{2}} = [e^{t^2}]^{-\frac{1}{2}} = e^{-\frac{1}{2}t^2}$$

$$\therefore \quad \lim_{n \to \infty} g(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} , -\infty < t < \infty$$

ان خاصية التقارب هذه تعني عملياً امكانية استخدام جداول التوزيع الطبيعي كبديل لجداول توزيع t عندما n كبيرة $(n > 30)$. وبملاحظة جدول توزيع t نلاحظ ان قيم هذا المتغير النظرية المقابلة لدرجة حرية $(n \to \infty)$ (السطر الاخير من الجدول) ماهي الا قيم Z لتوزيع $N(0,1)$ ازاء الاحتمال المقابل لكل قيمة من هذه القيم . فمثلاً لتوزيع t بدرجة حرية 120 فان :

$$P_r(t \leq 1{\cdot}98) \simeq P_r(z \leq 1{\cdot}96) = 0{\cdot}975$$

٩ ـ ٢ ـ ٧ : توزيع المؤشر الاحصائي $\frac{\sqrt{n}\,(\bar{X} - \mu)}{S}$

افرض ان $X_1, X_2, \ldots, X_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$. وبفرض ان σ^2 مجهولة وان $S^2, \bar{X}$ يمثلان على التوالي الوسط والتباين لقياسات هذه العينة . فاذا كان $\bar{X}$ مستقل عن S^2 , $ES^2 = \sigma^2$ عندئذٍ فان

$$t = \frac{\sqrt{n}\,(\bar{X} - \mu)}{S} \sim t_{(n-1)}$$

البرهان : سبق وان برهنا ان $\bar{X}$ مستقل عن S^2 وان $ES^2 = \sigma^2$ بحيث ان

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$ الان حيث ان $X \sim N(\mu, \sigma^2)$ فان

$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. وبذلك فان $Z_1 = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$. كذلك

$Z_2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$. وبفرض ان Z_1 مستقل عن Z_2 فان :

$$t = \frac{Z_1}{\sqrt{Z_2 / n - 1}} \sim t_{(n-1)}$$

اي ان :

$$t = \frac{\dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}}}{\sqrt{\dfrac{(n-1) S^2}{(n-1) \sigma^2}}} = \frac{\sqrt{n}\,(\bar{X} - \mu)}{S} \sim t_{(n-1)}$$

ان المؤشر الاحصائي اعلاه يمثل معيار لاختبار فرضية بشأن متوسط مجتمع طبيعي μ عندما تكون σ^2 مجهولة وان n صغيرة . اما اذا كانت n كبيرة فان $S^2 \xrightarrow[n \to \infty]{} \sigma^2$ وعندئذ فان

$$\lim_{n \to \infty} \frac{\sqrt{n}\,(\bar{X} - \mu)}{S} = \frac{\sqrt{n}\,(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$$

٩ ـ ٢ ـ ٨ : استخدامات توزيع t

ان لتوزيع t استخدامات كثيرة في تطبيقات النظرية الاحصائية نذكر بعضها ادناه دون الدخول في تفاصيلها .

أ ـ اختبار متوسط مجتمع طبيعي تباينه مجهول .
ب ـ اختبار الفرق بين متوسطي مجتمعين طبيعين مجهولي التباين .
جـ ـ اختبار معنوية معامل الارتباط.
د ـ اختبار معنوية معاملات الانحدار في نموذج انحدار خطي متعدد المتغيرات .
هـ ـ اختبار معنوية معامل الارتباط الجزئي .
و ـ تكوين حدود ثقة لمتوسط مجتمع طبيعي مجهول التباين .

مثال : افرض ان t متغير عشوائي وان $g(t) = C\left[1 + \frac{1}{5} t^2\right]^{-3}$ جد قيمة الثابت C بحيث ان t يتوزع كتوزيع t

الحل :

بمقارنة هذه الدالة مع الشكل العام لتوزيع t بـ n درجة حرية نلاحظ ان $n = 5$. عليه فان قيمة C ماهي الا :

$$C = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma\left(\frac{n}{2}\right)} = \frac{\Gamma\ (3)}{\sqrt{5\pi}\ \Gamma\left(\frac{5}{2}\right)} = 0{\cdot}37961$$

تمارين عن توزيع t

---

٩ ــ ٨ : اذا علمت ان $t \sim t_{(10)}$. يطلب اجراء مايلي :

أ ــ اكتب الدالة الاحتمالية الى t ثم ارسم مخطط هذه الدالة وقارنه مع مخطط $N(0, 1)$

ب ــ جد التباين ومعامل التفلطح لهذا التوزيع .

جـ ــ جد $P_r(t \leq 1)$, $P_r(t > 0{\cdot}7)$, $P_r(0{\cdot}5 < t < 1)$.

٩ ــ ٩ : افرض ان t يتوزع كتوزيع t بدرجة حرية واحدة . بين ان $F(t) = 0{\cdot}5 + \pi^{-1}\tan^{-1}t$.

٩ ــ ١٠ : افرض ان $t \sim t_{(n)}$. برهن ان $E|t| = \sqrt{n}\ \Gamma\left(\frac{n-1}{2}\right)\left[\sqrt{\pi}\ \Gamma\left(\frac{n}{2}\right)\right]^{-1}$. (ملاحظة : استفد من خاصية التماثل في هذا التوزيع)

٩ ــ ١١ : لتكن $X_1, X_2, \ldots, X_n$ تمثل قياسات عينة عشوائية من $N(\mu_x, \sigma_x^2)$ وان $Y_1, Y_2, \ldots, Y_m$ تمثل قياسات عينة عشوائية من $N(\mu_y, \sigma_y^2)$

وبفرض ان المجتمعان مستقلان وان $S_x^2, \bar{X}$ يمثلان على التوالي متوسط وتباين العينة الاولى وان $S_y^2, \bar{Y}$ متوسط وتباين العينة الثانية . وان متوسط العينة مستقل عن تباينها . برهن ان

$$t = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\left[\frac{(n-1) S_x^2 + (m-1) S_y^2}{n + m - 2} \left(\frac{1}{n} + \frac{1}{m} \right) \right]^{\frac{1}{2}}}$$

$$\sim t_{(n+m-2)}$$

٩ ـ ١٢ : اذا علمت ان $g(t) = c[3 + t^2]^{-2}$. جد قيمه c بحيث ان t يتوزع كتوزيع t .

٩ ـ ١٣ : اذا علمت ان $Z \sim N(0, 1)$ وان $X_1 \sim \chi^2_{(3)}$, $X_2 \sim \chi^2_{(4)}$, $X_3 \sim \chi^2_{(6)}$ وان هذه المتغيرات مستقلة تصادفياً . جد التوزيع الاحتمالي الى :

$$g_1 = \frac{\sqrt{3}\ Z}{\sqrt{X_1}}, \quad g_2 = \frac{\sqrt{7}\ Z}{\sqrt{X_1 + X_2}}, \quad g_3 = \frac{\sqrt{13}\ Z}{\sqrt{X_1 + X_2 + X_3}}$$

٩ ـ ١٤ : اذا علمت ان t يتوزع كتوزيع t بدرجة حرية واحدة . استخدم معطيات السؤال (٩ ـ ٩) لايجاد قيمة الربيع الاول والثالث . ماهي قيمة الربيع الثاني .

٩ ـ ٣ : توزيع F   F- distribution

يعتبر العالم R. A. Fisher اول من وضع اسس هذا التوزيع ونشرها عام ١٩٢٥ من خلال ايجاد علاقة بين هذا التوزيع والتوزيع الطبيعي ، في حين قام G. W. Snedecor باجراء تعديلات على انجازات فيشر ونشرها عام ١٩٣٤ وابقى انجازاته هذه مقرونة باسم فيشر من خلال تسميته لهذا التوزيع بتوزيع F. ويعد توزيع F احد التوزيعات المهمة جداً في الكثير من التطبيقات الاحصائية وخصوصاً في موضوع تحليل التباين Analysis of variance . وفيما يلي عرض لطريقة Snedecor في اشتقاق دالة هذا التوزيع .

٩ ـ ٣ ـ ١ : اشتقاق دالة توزيع F .

افرض ان $X_1 \sim \chi^2_{(n_1)}$ مستقل عن $X_2 \sim \chi^2_{(n_2)}$. عندئذٍ فان النسبة $f = \dfrac{x_1 / n_1}{x_2 / n_2}$ سوف تتوزع كتوزيع F بدرجتي حرية n_1 , n_2 .

البرهان :

ان التوزيع المشترك للمتغيرين X_1 , X_2 يتمثل بدالة الكثافة الاحتمالية المشتركة الناتجة من حاصل ضرب $f(x_1)$ بـ $f(x_2)$. اي ان

$$f(x_1 , x_2) = \frac{1}{\Gamma\left(\frac{n_1}{2}\right) . 2^{\frac{n_1}{2}}} \, x_1^{\frac{n_1}{2} - 1} . e^{-\frac{x_1}{2}} . \frac{1}{\Gamma\left(\frac{n_2}{2}\right) . 2^{\frac{n_2}{2}}} \, x_2^{\frac{n_2}{2} - 1} . e^{-\frac{x_2}{2}}$$

$$= \frac{1}{\Gamma\left(\frac{n_1}{2}\right) . \Gamma\left(\frac{n_2}{2}\right) . 2^{(n_1 + n_2)/2}} \, x_1^{\frac{n_1}{2} - 1} . x_2^{\frac{n_2}{2} - 1} . e^{-\frac{1}{2}(x_1 + x_2)}$$

الان

$$f = \frac{n_2}{n_1} \cdot \frac{x_1}{x_2} \rightarrow f > 0 , x_1 > 0 , x_2 > 0$$

وبفرض ان $y = x_2$ فان $x_2 = y$ وان

$$f = \frac{n_2}{n_1} \cdot \frac{x_1}{y}$$

فاذن $x_1 = \frac{n_1}{n_2} \cdot fy$ وان معامل التحويل (جاكوبيان) J هو :

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial f} & \frac{\partial x_1}{\partial y} \\ \frac{\partial x_2}{\partial f} & \frac{\partial x_2}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{n_1}{n_2} y & \frac{n_1}{n_2} f \\ 0 & 1 \end{vmatrix} = \frac{n_1}{n_2} \cdot y$$

$$\therefore | J | = \frac{n_1}{n_2} \cdot y$$

عليه فان

$$g (f , y) = f (x_1 , x_2) \Big]_{\substack{x_1 = \frac{n_1}{n_2} fy \\ x_2 = y}} \cdot | J |$$

$$= K \left(\frac{n_1}{n_2} fy \right)^{\frac{n_1}{2} - 1} \cdot y^{\frac{n_2}{2} - 1} \cdot e^{- \frac{1}{2} \left(\frac{n_1}{n_2} fy + y \right)} \cdot \frac{n_1}{n_2} y \; ;$$

$$K = \frac{1}{\Gamma \left(\frac{n_1}{2} \right) \Gamma \left(\frac{n_2}{2} \right) . 2^{(n_1 + n_2)/2}}$$

$$\therefore g(f,y) = K\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} . f^{\frac{n_1}{2}-1} . y^{\frac{n_1+n_2}{2}-1} . e^{-\frac{1}{2}y\left(\frac{n_1}{n_2}f+1\right)}$$

فاذن

$$g(f) = K\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} . f^{\frac{n_1}{2}-1} \int_0^{\infty} y^{\frac{n_1+n_2}{2}-1} e^{-\frac{1}{2}y\left(\frac{n_1}{n_2}f+1\right)} dy$$

لكن التكامل في الصيغة الاخيرة يمثل تكامل كاما بالمعلمتين

$\alpha = \frac{n_1+n_2}{2}$, $\beta = 2\left(\frac{n_1}{n_2}f+1\right)^{-1}$ وان قيمة هذا التكامل هي

$$\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{2}{\frac{n_1}{n_2}f+1}\right)^{\frac{n_1+n_2}{2}}$$

عليه فان

$$g(f) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right).\Gamma\left(\frac{n_2}{2}\right).2^{\frac{n_1+n_2}{2}}} . \frac{2^{\frac{n_1+n_2}{2}}}{\left(1+\frac{n_1}{n_2}f\right)^{\frac{n_1+n_2}{2}}} . f^{\frac{n_1}{2}-1}$$

$$\therefore g(f) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} . \frac{f^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2}f\right)^{\frac{n_1+n_2}{2}}}, f>0$$

ان الدالة $g(f)$ تسمى دالة توزيع F بـ n_2 , n_1 درجة حرية . وبالرموز فان $f \sim F(n_1, n_2)$. ان n_1 تعني درجات الحرية للمتغير المعرف في بسط النسبة f في حين ان n_2 تعني درجات الحرية للمتغير المعرف في مقامها .

٩ ـ ٣ ـ ٢ : الدالة التوزيعية لتوزيع F

ان جداول توزيع F (لاحظ الجدول ٨ ملحق ب) الخاصة بقيم F النظرية تم حسابها بالاعتماد على دالة بيتا غير التامة incomplete beta function حيث ان هذا الاسلوب يعطي ادق قيم نظرية لهذا التوزيع . فعلى فرض ان $f \sim F(n_1, n_2)$ وان $G(f)$ تمثل الدالة التوزيعية لهذا التوزيع فان :

$$G(f_0) = P_r(f \le f_0)$$

حيث ان f_0 تمثل قيمة من قيم f المعرفة في الفترة $(0, \infty)$ فاذن

$$G(f_0) = \int_0^{f_0} g(f)\,df$$

$$= \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right).\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}\int_0^{f_0}\frac{f^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2}f\right)^{\frac{n_1+n_2}{2}}}\,df\,.$$

الان بفرض ان $y = \dfrac{n_1 f}{n_2 + n_1 f}$ او ان $y = \dfrac{1}{\frac{n_2}{n_1 f} + 1}$ وذلك يعني انه عندما $f = 0$ فان $y = 0$ وعندما $f \to \infty$ فان $y \to 1$. وهذا يعني ان $0 < y < 1$. كذلك فان $f = \dfrac{n_2}{n_1}\left(\dfrac{y}{1-y}\right)$ فاذن $df = \dfrac{n_2}{n_1} \cdot \dfrac{dy}{(1-y)^2}$.

عليه فان :

$$G(f_0) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}$$

$$\int_0^{y_0}\frac{\left[\frac{n_2}{n_1}\left(\frac{y}{1-y}\right)\right]^{\frac{n_1}{2}-1}}{\left(1+\frac{y}{1-y}\right)^{\frac{n_1+n_2}{2}}}\cdot\left(\frac{n_2}{n_1}\right).\frac{dy}{(1-y)^2}\ ;$$

$$y_0 = \frac{n_1 f_0}{n_2 + n_1 f_0}$$

$$= \int_0^{y_0}\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right).\Gamma\left(\frac{n_2}{2}\right)}\cdot y^{\frac{n_1}{2}-1}.(1-y)^{\frac{n_2}{2}-1}dy$$

ان التكامل الاخير يمثل تعريف الدالة التوزيعية لتوزيع بيتا بالمعلمتين $\alpha = \frac{n_1}{2}$ و $\beta = \frac{n_2}{2}$ وهذا يعني امكانية استخدام هذه الدالة (جداول توزيع بيتا) في تكوين جداول توزيع F . فمثلاً اذا كان $f \sim F(4, 4)$ وتطلب الامر حساب $P_r(f \leq 6{\cdot}39)$ فان ذلك يتم وفق الآتي :

$$y_0 = \frac{n_1 f_0}{n_2 + n_1 f_0} = \frac{4(6{\cdot}39)}{4 + 4(6{\cdot}39)} = \frac{25{\cdot}56}{29{\cdot}56} = 0{\cdot}864682$$

فاذن

$$P_r(f \le 6{\cdot}39) = G(6{\cdot}39) = \int_0^{0{\cdot}864682} \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} y(1-y)dy$$

$$= 6\left[\left[\frac{y^2}{2}\right]_0^{y_0} - \left[\frac{y^3}{3}\right]_0^{y_0} \right] = 6[0{\cdot}3738374 - 0{\cdot}2155003]$$

$$= 0{\cdot}9500226 \simeq 0{\cdot}95$$

وعلى هذا الاساس تم بناء جداول توزيع F بحيث ان القيم الموضحة في خلايا هذا الجدول تمثل قيم f النظرية عند درجتي حرية n_1 للبسط و n_2 للمقام . وللاغراض التطبيقية فانه غالباً مايصار الى عرض نوعين من هذه الجداول الاول خاص بقيم f التي تعطي احتمال متراكم قدره 0·95 (او لمستوى معنوية 0·05) والثاني خاص بقيم f التي تعطي احتمال متراكم قدره 0·99 (او لمستوى معنوية 0·01) اي مانعنيه : $P_r(f \le f_0) = 0{\cdot}95 , P_r(f \ge f_0) = 0{\cdot}05 = \alpha$

وان

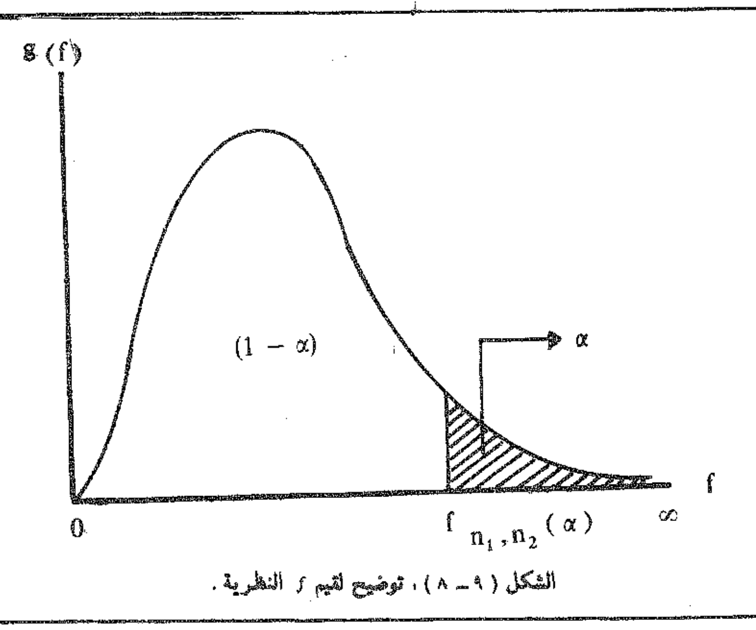
$$P_r(f \le f_0) = 0{\cdot}99 , P_r(f \ge f_0) = 0{\cdot}01$$

وغالباً مايرمز لقيم f النظرية بالشكل $f_{n_1,n_2}(\alpha)$ اي قيمة f النظرية بدرجتي حرية n_2 , n_1 ومستوى معنوية α . وسوف نلاحظ في فقرة لاحقة ان توزيع F هو ذا التواء موجب دائماً . والشكل (٩ ــ ٨) يبين عملية حساب القيم النظرية لهذا التوزيع .

علماً ان هنالك طرقاً اخرى مقترحة لحساب الدالة التوزيعية لتوزيع F كتلك المقترحة من قبل Zinger* عام ١٩٦٤ و Wishart عام ١٩٥٧ وغيرهم الا انها طرقاً تقريبية ليست افضل من الاسلوب الذي استعرضناه .

---

* ، للمزيد من التفاصيل انظر المصدر (9) .



الشكل (٩ ـ ٨) : توضيح لقيم f النظرية .

٩ ـ ٣ ـ ٣ : عزوم توزيع F

نظراً لصعوبة صياغة دالة مولده لعزوم توزيع F بشكل مألوف فاننا سوف نستعيض عنها من خلال ايجاد صيغة للعزم ذا المرتبة r حول نقطة الاصل طالما ان الهدف واحد في كلا الحالتين وهو استنتاج عزوم التوزيع وبشكل خاص الوسط والتباين . فاذا كان $f \sim F(n_1, n_2)$ فان :

$$Ef^r = E\left(\frac{n_2}{n_1} \cdot \frac{X_1}{X_2}\right)^r = \left(\frac{n_2}{n_1}\right)^r . EX_1^r X_2^{-r}$$

وحيث ان $X_1 \sim \chi^2_{(n_1)}$ مستقل عن $X_2 \sim \chi^2_{(n_2)}$ فاذن :

$$Ef^r = \left(\frac{n_2}{n_1}\right)^r EX_1^r . EX_2^{-r}$$

وحسب ماهو موضح في الفقرة (٩ ـ ١) لدى دراستنا لتوزيع مربع كاي فان :

$$EX_1^r = 2^r \cdot \frac{\Gamma\left(\frac{n_1}{2}+r\right)}{\Gamma\left(\frac{n_1}{2}\right)}, EX_2^{-r} = 2^{-r} \frac{\Gamma\left(\frac{n_2}{2}-r\right)}{\Gamma\left(\frac{n_2}{2}\right)}$$

عليه فان :

$$Ef^r = \left(\frac{n_2}{n_1}\right)^r \frac{\Gamma\left(\frac{n_1}{2}+r\right).\Gamma\left(\frac{n_2}{2}-r\right)}{\Gamma\left(\frac{n_1}{2}\right).\Gamma\left(\frac{n_2}{2}\right)}, \quad \frac{n_2}{2} > r, r = 1,2,\ldots$$

وبشكل خاص عندما $r = 1$ فان متوسط توزيع F هو

$$\mu_f = Ef = \left(\frac{n_2}{n_1}\right) \cdot \frac{\Gamma\left(\frac{n_1}{2}+1\right)\Gamma\left(\frac{n_2}{2}-1\right)}{\Gamma\left(\frac{n_1}{2}\right).\Gamma\left(\frac{n_2}{2}\right)} = \frac{n_2}{n_2-2}, n_2 > 2$$

ويتضح من ذلك ان متوسط توزيع F مستقل عن درجات حرية البسط n_1.
وعندما $r = 2$ فان :

$$Ef^2 = \frac{n_2^2(n_1+2)}{n_1(n_2-2)(n_2-4)}; n_2 > 4$$

عليه فان تباين توزيع F هو

$$V(f) = Ef^2 - \mu_f^2 = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2.(n_2-4)}; n_2 > 4$$

$$= 2\mu_f^2 \cdot \frac{n_1+n_2-2}{n_1(n_2-4)}$$

٩ ـ ٣ ـ ٤ : المنوال ونقاط الانقلاب في دالة توزيع F

كما هو معلوم فان المنوال ناتج من حل المعادلة التفاضلية $g'(f) = 0$ بشرط ان $g''(f) < 0$. فاذا رمزنا لثابت دالة توزيع F بـ K فان

$$\log g(f) = \log k + \left(\frac{n_1}{2} - 1\right)\log f - \left(\frac{n_1 + n_2}{2}\right)\log\left(1 + \frac{n_1}{n_2} f\right)$$

وبايجاد مشتقة الطرفين نسبة الى f نحصل على :

$$\frac{g'(f)}{g(f)} = \frac{n_1 - 2}{2f} - \left(\frac{n_1 + n_2}{2}\right) \cdot \frac{n_1}{n_2 + n_1 f}$$

$$\therefore g'(f) = g(f)\left[\frac{n_1 - 2}{2f} - \frac{n_1(n_1 + n_2)}{2(n_2 + n_1 f)}\right]$$

وبجعل $g'(f) = 0$ وحل المعادلة نسبة الى f نحصل على :

$$f = \frac{n_2(n_1 - 2)}{n_1(n_2 + 2)} = \frac{1 - 2/n_1}{1 + 2/n_2} < 1$$

ونترك للقاريء البيان ان $g''(f) < 0$ عندما $f = \frac{n_2(n_1 - 2)}{n_1(n_2 + 2)}$. وهذا يعني

ان قيمة المنوال في توزيع F هي $f = \frac{n_2(n_1 - 2)}{n_1(n_2 + 2)}$ وهي دائماً اقل من واحد بشرط ان $n_1 \geq 2$.

كذلك فان لمنحنى دالة هذا التوزيع نقطتي انقلاب عندما $n_1 > 4$ تقعان على بُعد متساوي الى يمين ويسار المنوال ناتجتين من حل المعادلة التفاضلية $g''(f) = 0$ بشرط ان $g'''(f) \neq 0$ الا انه من الصعوبة ايجاد صيغة لكل منهما . الا انه امكن البرهنة على ان مجموع هاتين النقطتين هو $\text{Mode} \cdot 2 = \frac{2n_2(n_1 - 2)}{n_1(n_2 + 2)}$

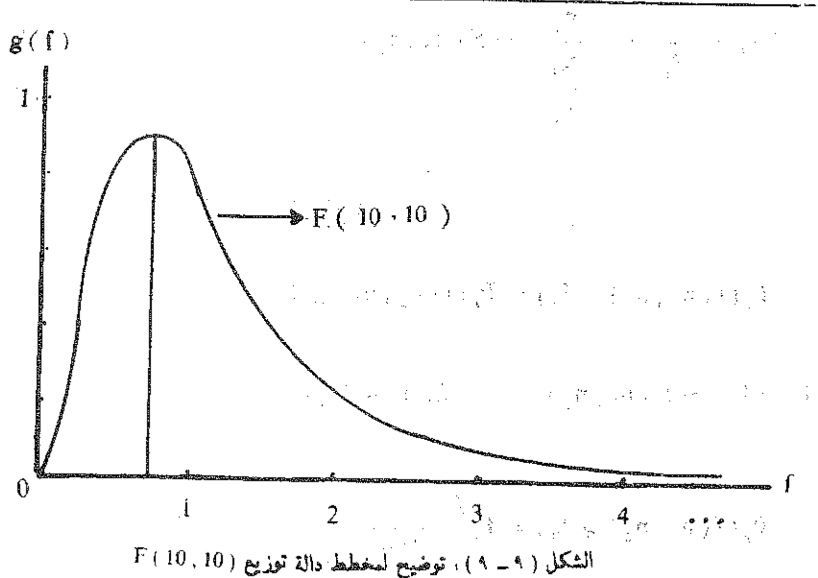
٩ ـ ٢ ـ ٥ : الالتواء في توزيع F

لاحظنا مما تقدم ان $Mode = \frac{n_2 (n_1 - 2)}{n_1 (n_2 + 2)} < 1$ وان $Mean = \frac{n_2}{n_2 - 2} > 1$

وهذا يعني ان معامل التواء التوزيع وفق صيغة كارل بيرسون هي :

$$S_k = \frac{mean - Mode}{\sqrt{V(f)}} > 0$$

وهذا يعني ان منحنى دالة هذا التوزيع ملتو التواء موجب دائماً وتزداد شدة الالتواء بانخفاض قيمة n_1 بفرض ثبات قيمة n_2 . والشكل (٩ ـ ٩) يوضح مخطط لدالة توزيع F (10 , 10)



الشكل (٩ ـ ٩) : توضيح لمخطط دالة توزيع F (10 , 10)

٩ ـ ٣ ـ ٦ : خاصية الانعكاس في توزيع F .

ليكن $f_1 \sim F(n_1, n_2)$ عندئذ فان $f_2 = \frac{1}{f_1} \sim F(n_2, n_1)$

البرهان : من تعريف توزيع F فان

$$f_1 = \frac{\frac{X_1}{n_1}}{\frac{X_2}{n_2}} \sim F(n_1, n_2) ; X_1 \sim \chi^2_{(n_1)}, X_2 \sim \chi^2_{(n_2)}$$

$$\therefore f_2 = \frac{1}{f_1} = \frac{\frac{X_2}{n_2}}{\frac{X_1}{n_1}} \sim F(n_2, n_1)$$

واستناداً لهذه الخاصية يمكن البيان ان

$$P_r(f(n_1, n_2) \geq f_0) = P_r(f(n_2, n_1) \leq f_0^{-1}$$

لاحظنا انه عندما $f_1 \sim F(n_1, n_2)$ فان $f_2 = f_1^{-1} \sim F(n_2, n_1)$
فاذن

$$P_r(f(n_1, n_2) \geq f_0) = P_r\left(\frac{1}{f(n_1, n_2)} \leq \frac{1}{f_0} \right)$$

لكن $\frac{1}{f(n_1, n_2)}$ ما هو الا متغير عشوائي يتوزع كتوزيع $F(n_2, n_1)$. فاذن

$$P_r(f(n_1, n_2) \geq f_0) = P_r(f(n_2, n_1) \leq f_0^{-1})$$

وبشكل خاص اذا كان

$$P_r(f(n_1,n_2) \geq f_0) = \alpha \rightarrow P_r(f(n_1,n_2) \leq f_0) = 1 - \alpha, 0 < \alpha < 1$$

فان

$$P_r(f(n_2, n_1) \leq f_0^{-1}) = \alpha$$

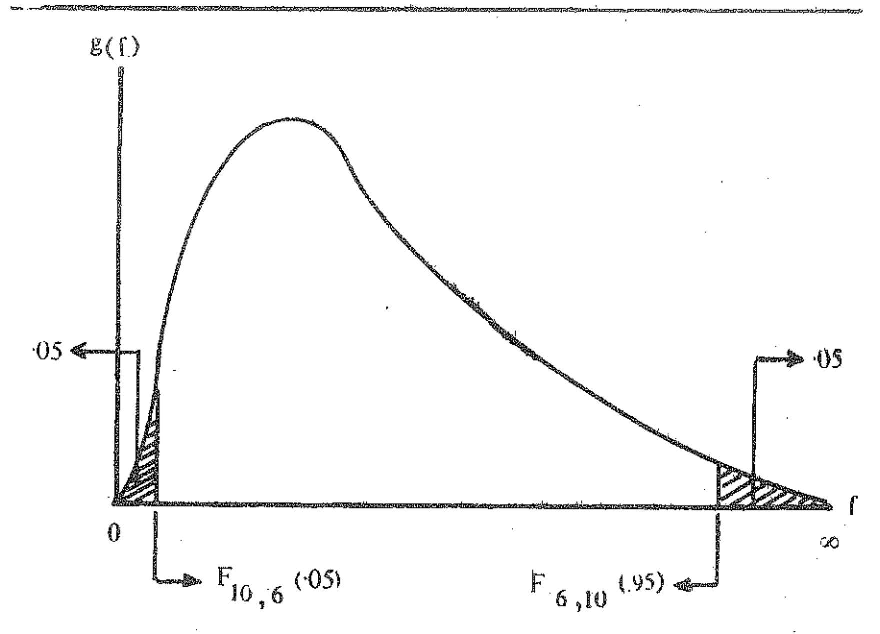
وان

$$1 - P_r(f(n_2, n_1) \leq f_0^{-1}) = 1 - \alpha$$

أي ان

$$P_r(f(n_2, n_1) \geq f_0^{-1}) = 1 - \alpha$$

وهذا يعني ان قيمة $f(n_1, n_2)$ التي تعطي احتمالاً متراكماً قدره $1 - \alpha$ هي معكوس قيمة $f(n_2, n_1)$ التي تعطي احتمالاً متراكماً قدره α والعكس صحيح. على سبيل المثال ولتوزيع $F(6, 10)$ ومن جداول هذا التوزيع نلاحظ ان قيمة $F_{6,10}(0.95) = 3.22$ وهذا يعني ان $F_{10,6}(0.05) = \frac{1}{3.22}$. وكما هو موضح في الشكل (٩ ــ ١٠) :



الشكل (٩ ــ ١٠) ، توضيح لقيم f ومعكوسها .

ويتضح من الشكل (٩ ــ ١٠) ان

$$P_r (F_{10,6} (0{\cdot}05) < f < F_{6,10} (0{\cdot}95)) = 0{\cdot}90$$

٩ ــ ٣ ــ ٧ : توزيع النسبة بين تبايني عينتين مستقلتين :

افرض ان S_1^2 يمثل تباين عينة عشوائية قوامها n_1 مسحوبة من مجتمع طبيعي تباينه مجهول وليكن σ_1^2 . وافرض ان S_2^2 يمثل تباين عينة اخرى قوامها n_2 مسحوبة من مجتمع طبيعي تباينه مجهول وليكن σ_2^2 . وبفرض ان العينتين مستقلتان عندئذٍ فان المؤشر الاحصائي

$$f = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F (n_1 - 1 , n_2 - 1)$$

البرهان : من المعلوم ان

$$Z_1 = \frac{(n_1 - 1) S_1^2}{\sigma_1^2} \sim \chi^2 (n_1 - 1)$$

وان

$$Z_2 = \frac{(n_2 - 1) S_2^2}{\sigma_2^2} \sim \chi^2 (n_2 - 1)$$

وحيث العينتين مستقلتان فان :

$$f = \frac{\dfrac{Z_1}{n_1 - 1}}{\dfrac{Z_2}{n_2 - 1}} \sim F (n_1 - 1 , n_2 - 1)$$

وبالتعويض عن Z_2 , Z_1 بما يساويهما نجد ان

$$f = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F (n_1 - 1 , n_2 - 1).$$

وبشكل خاص اذا كنا بصدد اختبار الفرض القائل بأن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ فان

$$f = \frac{S_1^2}{S_2^2} \sim F (n_1 - 1 , n_2 - 1)$$

وهذا يمثل معيار اختبار فرضية تجانس تبايني مجتمعين طبيعين ويمثل الاساس في موضوع تحليل التباين . وغالباً مايتم واثناء التطبيق وضع التباين الاكبر في البسط والتباين الاصغر في المقام فللحالة السابقة افترضنا ان $S_1^2 > S_2^2$. ان السبب في هذا الاجراء يعود الى ان القيم النظرية الموضحة في جداول توزيع F هي دائماً اكبر من الواحد وفي حالة وضع التباين الاصغر في البسط فأن $f < 1$ وفي هذه الحالة (ومن الناحية العملية) لايمكن اجراء المقارنة (عند اختبار فرضية معينة مثلاً) بين قيمة f المستخرجة وقيمة f النظرية .

٩ ـ ٣ ـ ٨ : استخدامات توزيع F .

ان لتوزيع F استخدامات عديدة في تطبيقات النظرية الاحصائية نذكر منها الاتي دون الدخول في تفاصيلها .

أ ـ اختبار F لتجانس تبايني عينتين مستقلتين .

ب ـ اختبار معنوية معامل الارتباط المتعدد .

جـ ـ اختبار معولية نموذج انحدار خطي متعدد .

د ـ اختبار F لتجانس عدة تقديرات مستقلة لتباين مجتمع طبيعي .

هـ ـ الاختبارات الخاصة بتحليل التباين والتباين المشترك .

تمارين عن توزيع F

٩ ـ ١٥ : افرض ان $f \sim F(6,6)$. يطلب اجراء ما يلي :

أ ـ جد دالة الكثافة الاحتمالية الى f ثم ارسم مخطط هذه الدالة .

ب ـ جد $P_r(f \geq 4.28), P_r(f \leq 3)$

جـ ـ جد الوسط والتباين والمنوال وقيمة معامل التواء التوزيع .

٩ ـ ١٦ : أذا كان $f \sim F(n_1, n_2)$ برهن ان $Y = \left(1 + \frac{n_1}{n_2} f\right)^{-1}$ يتوزع

كتوزيع بيتا بالمعلمتين $\beta = n_1/2, \alpha = n_2/2$.

٩ ـ ١٧ : اذا كان $f \sim F(n, m)$ برهن باستخدام اسلوب التحويلات ان

$f^{-1} \sim F(m,n)$

٩ ـ ١٨ : اذا علمت ان $f \sim F(2, n)$. برهن ان

$$P_r(f \geq f_0) = \left(1 + \frac{2}{n} f_0\right)^{-\frac{n}{2}}$$

٩ ـ ١٩ : اذا كان $f_1 \sim F(m, n)$ وان $f_2 \sim F(n, m)$. برهن ولاي عدد

موجب مثل a ان $P_r(f_1 \leq a) + P_r\left(f_2 \leq \frac{1}{a}\right) = 1$

٩ ـ ٢٠ : اذا كان $f \sim F(2,2)$. بين ولاي عدد موجب مثل f_0 ان

$P_r(f \leq f_0) = f_0 / (f_0 + 1)$

٩ ــ ٤ : العلاقة بين توزيعات المعاينة

اختصت الفقرات السابقة لهذه الفقرة باستعراض لاهم توزيعات المعاينة الشائعة الاستخدام في التطبيقات الاحصائية . في هذه الفقرة سوف ندرس بعض العلاقات التي تربط هذه التوزيعات مع بعضها وهي :

٩ ــ ٤ ــ ١ : العلاقة بين توزيع t وتوزيع F

تنص هذه العلاقة بمايلي : اذا كان $t \sim t_{(n)}$ فان $f = t^2 \sim F(1, n)$

البرهان :

بما ان $f = t^2$, $-\infty < t < \infty$ فان $0 < f < \infty$ وان $t = \pm \sqrt{f}$

الان $t = \sqrt{f}$ لقيم $t > 0$ وهذا يعني ان $dt = \frac{1}{2} f^{-\frac{1}{2}} df$ وان

(بسبب خاصية تماثل توزيع t) $2 \int_0^\infty g(t)\, dt = 1$

$$\Longrightarrow 2 \int_0^\infty \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt = 1$$

$$\Longrightarrow 2 \int_0^\infty \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{f}{n}\right)^{-\frac{n+1}{2}} \frac{1}{2} f^{-\frac{1}{2}}\ df = 1$$

$$\Longrightarrow \int_0^\infty \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} \left(\frac{1}{n}\right)^{\frac{1}{2}} \frac{f^{-\frac{1}{2}}}{\left(1+\frac{f}{n}\right)^{\frac{n+1}{2}}} df = 1 ,$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$\Longrightarrow \int_0^\infty g(f ; n_1 = 1, n_2 = n) df = 1$$

$$\therefore g(f) = \frac{\Gamma\left(\frac{n+1}{2}\right)\left(\frac{1}{n}\right)^{-\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{f^{-\frac{1}{2}}}{\left(1+\frac{f}{n}\right)^{\frac{n+1}{2}}} ;$$

$$0 < f < \infty$$

وهذا يعني ان $f = t^2 \sim F(1, n)$

ان هذه العلاقة تسمح لنا استخدام جداول توزيع t كبديل لجداول توزيع F في حالة كون درجة حرية البسط مساوية للواحد. هذه العلاقة من الناحية العملية هي

$$0 < \alpha < 1 , F_{1,n\,(1-\alpha)} = t^2_{n\,(1-\alpha/2)}$$

فمثلاً عندما $\alpha = 0.05 , n = 10$ فانه من جداول توزيع F عند درجتي حرية $(1, 10)$ باحتمال متراكم $1 - \alpha = 0.95$ نجد ان $F_{1,10}(0.95) = 4.96$ وهذه تمثل في ذات الوقت مربع $t_{10}(0.975) = 2.228$ اي $4.963984 \simeq 4.96$ وبشكل خاص فان قيم العمود الاول من جداول توزيع F باحتمال متراكم 0.95 ما هي الا مربع قيم العمود الخامس في جداول توزيع t باحتمال متراكم 0.975.

٩ ـ ٤ ـ ٢ : العلاق بين توزيع F وتوزيع χ^2.

تنص هذه العلاقة بما يلي : اذا كان $f \sim F(n_1, n_2)$ وان $X = n_1 F$ فان التوزيع التقاربي الى X عندما $n_2 \to \infty$ هو توزيع مربع كاي بـ n_1 درجة حرية.

البرهان : ان

$$g(f) = \frac{\Gamma\left(\frac{n_1 + n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) \cdot \Gamma\left(\frac{n_2}{2}\right)} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{f^{\frac{n_1}{2} - 1}}{\left(1 + \frac{n_1}{n_2} f\right)^{\frac{n_1 + n_2}{2}}} \quad , f > 0$$

$$= \quad A \qquad B \qquad C$$

الان

$$\lim_{n_2 \to \infty} g(f) = \lim_{n_2 \to \infty} A \cdot \lim_{n_2 \to \infty} B \cdot \lim_{n_2 \to \infty} C$$

$$\therefore \lim_{n_2 \to \infty} A = \frac{1}{\Gamma\left(\frac{n_1}{2}\right)} \lim_{n_2 \to \infty} \frac{\Gamma\left(\frac{n_1}{2} + \frac{n_2}{2}\right)}{\Gamma\left(\frac{n_2}{2}\right)} = \frac{\left(\frac{n_2}{2}\right)^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right)}$$

$$n_2 \to \infty$$

حيث ان وبشكل عام $\lim_{n \to \infty} \frac{\Gamma(n + k)}{\Gamma(n)} = n^k$ (لاحظ برهان ذلك في الفقرة (٩ ـ ٢ ـ ٦)

وان

$$\lim_{n_2 \to \infty} B = \lim_{n_2 \to \infty} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} = \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \quad , n_2 \to \infty$$

وان

$$\lim_{n_2\to\infty} C = f^{\frac{n_1}{2}-1} \cdot \lim_{n_2\to\infty}\left(1+\frac{n_1}{n_2}f\right)^{-\frac{n_1}{2}} \cdot \lim_{n_2\to\infty}\left[\left(1+\frac{n_1}{n_2}f\right)^{n_2}\right]^{-\frac{1}{2}}$$

$$= f^{\frac{n_1}{2}-1} \cdot (1) \cdot e^{-\frac{1}{2}n_1 f}$$

$$\therefore \lim_{n_2\to\infty} g(f) = \lim_{n_2\to\infty} \frac{1}{\Gamma\left(\frac{n_1}{2}\right)} \cdot \left(\frac{n_2}{2}\right)^{\frac{n_1}{2}} \cdot \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \cdot f^{\frac{n_1}{2}-1} \cdot e^{-\frac{1}{2}n_1 f}$$

$$= \lim_{n_2\to\infty} \frac{n_1^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right) \cdot 2^{n_1/2}} \cdot f^{\frac{n_1}{2}-1} \cdot e^{-\frac{1}{2}n_1 f}$$

$$= \frac{n_1^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right) 2^{n_1/2}} \cdot f^{\frac{n_1}{2}-1} \cdot e^{-\frac{1}{2}n_1 f}$$

وحيث ان $0 < f < \infty , X = n_1 f$ فان $0 < x < \infty$. وان $f = \frac{X}{n_1}$ فان

فاذن $g(x) = \lim_{n_2\to x} g(f)\Big]_{f=\frac{x}{n_1}} \cdot \left|\frac{df}{dx}\right|$ عليه فان . $\frac{df}{dx} = \frac{1}{n_1}$

$$g(x) = \frac{n_1^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right) \cdot 2^{n_1/2}} \left(\frac{x}{n_1}\right)^{\frac{n_1}{2}-1} \cdot e^{-\frac{1}{2}x} \cdot \frac{1}{n_1}$$

$$\therefore g(x) = \frac{1}{\Gamma\left(\frac{n_1}{2}\right) . 2^{n_1/2}} . x^{\frac{n_1}{2} - 1} . e^{-\frac{x}{2}} \; ; x > 0$$

والدالة الاخيرة تمثل دالة توزيع مربع كاي بـ n_1 درجة حرية . فاذن

$$x = n_1 . F(n_1, n_2) \underset{n_2 \to \infty}{\sim} \chi^2_{(n_1)}$$

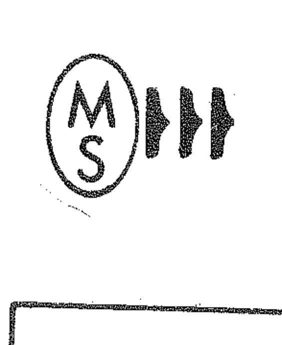
ان هذه العلاقة تسمح لنا استخدام جداول توزيع χ^2 كبديل لجداول توزيع F عندما تكون n_2 كبيرة (نظرياً $n_2 \to \infty$) اي ان

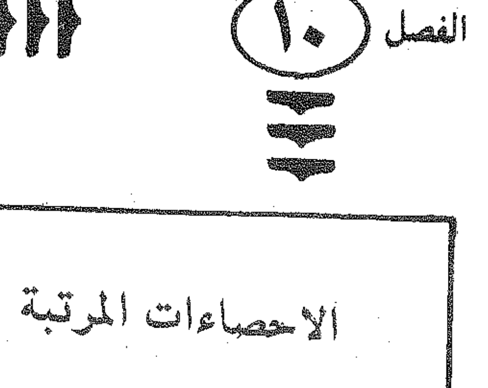
$$\chi^2_{(n_1)}(1 - \alpha) = n_1 . F_{(n_1, n_2 \to \infty)}(1 - \alpha), 0 < \alpha < 1$$

فمثلاً عندما $n_1 = 10$, $n_2 \to \infty$, $\alpha = 0.05$ فان قيمة F النظرية (من جداول توزيع F) هي 1.83 في حين ان القيمة المقابلة لها في جداول توزيع مربع كاي هي عند درجة حرية 10 باحتمال متراكم 0.95 هي 18.3 التي هي في الحقيقة تمثل $18.3 = 10(1.83) = n_1F$. وبشكل خاص فان قيم الصف الاخير من جداول توزيع F (المقابل الى $n_2 \to \infty$) باحتمال متراكم 0.95 ماهي الا حاصل قسمة قيم العمود الثالث في جداول توزيع χ^2 ، عند احتمال متراكم 0.95 ، على درجات الحرية المقابلة لها والعكس صحيح ايضاً وكذلك الحال بالنسبة لاي احتمال متراكم آخر (0.90 , 0.975 , 0.99 ...) . على سبيل المثال لو تطلب الامر حساب $F_{10, 1000}(0.99)$ فان هذه القيمة غير معرفة في جداول توزيع F. الا انه يمكن حسابها من جداول توزيع χ^2 وفق الاتي :

$$F_{10, 1000}(0.99) = \frac{\chi^2_{10}(0.99)}{10} = \frac{23.2093}{10} \simeq 2.32$$

٥٦٦





الفصل ١٠

الاحصاءات المرتبة

الفصل العاشر

الاحصاءات المرتبة Order statistics

سبق وان اشرنا في الفقرة (٨ ـ ١ ـ ٢) الى مفهوم المؤشر الاحصائي وذكرنا بانه دالة بدلالة قياسات عينة عشوائية خالية من اي مجهول . كذلك تعرفنا الى عزوم المؤشر الاحصائي (خصوصاً ما يتعلق الامر بالوسط الحسابي والتباين) كذلك توزيع المؤشر الاحصائي (لاحظ الفقرة ٨ ـ ١ ـ ٣) . في هذا الفصل سوف نتطرق الى مفهوم آخر يستند بالاساس لمفهوم العينة العشوائية وهو الاحصاءات المرتبة (او القيم المرتبة) وذلك من خلال استعراض لمفهوم الاحصاءات المرتبة وتوزيعاتها الاحتمالية . وكما لاحظنا في الفصل الثامن فان العزوم المستحصل عليها من العينة عبارة عن تقديرات تخص عزوم المجتمع كمعالم . فالوسط الحسابي $\bar{X}$ مثلاً يمثل العزم ذا المرتبة الاولى حول نقطة الاصل وهو في ذات الوقت تقدير الى متوسط المجتمع μ ، في حين ان الاحصاءات المرتبة عبارة عن تقديرات الى تجزئات quantiles تخص المجتمع كالوسيط والربيعات والعشيرات وغيرها من المقاييس التجزيئية .

١٠ ـ ١ : تعريف القيم المرتبة .

افرض ان $X_1, X_2, \ldots, X_n$ تمثل قياسات عينة عشوائية قوامها n مسحوبة من مجتمع بدالة كثافة احتمالية $a < x < b, f(x)$ حيث a وb عددان حقيقيان معرفان في حقل الاعداد الحقيقية . وافرض ان Y_1 تمثل اصغر قيمة من بين قيم هذه العينة وان Y_2 تمثل القيمة التالية من حيث الكبر وهكذا فان Y_n تمثل اكبر قيمة من بين قيم هذه العينة . وهذا يعني ان $Y_1 \leq Y_2 \leq Y_3 \leq \ldots \leq Y_n$ يمثل « ترتيب » قياسات العينة وفق ترتيب تصاعدي (وقد يكون الترتيب تنازلي) . هذا الترتيب يسمى الاحصاءات المرتبة (او القيم المرتبة) ، وعندئذٍ يقال ان Y_i يمثل القيمة (المؤشر الاحصائي) المرتب ذات التسلسل (الرتبة) i للعينة العشوائية $X_1, X_2, \ldots, X_n$. وحيث ان قياسات العينة العشوائية تعتبر بحكم المتغيرات العشوائية المستقلة ذات نفس التوزيع المعرف

بالدالة $f(x)$ فان الاحصاءات المرتبة وعلى الرغم من كونها قيم نفس العينة مرتبة بشكل تصاعدي (او تنازلي) الا انه لايمكن اعتبارها متغيرات مستقلة بسبب اعتماد القيمة اللاحقة في الترتيب على سابقتها .

١٠ ـ ٢ : التوزيع المشترك للاحصاءات المرتبة

ان التوزيع المشترك للاحصاءات المرتبة لقيم عينة عشوائية قوامها n مفردة قياساتها هي $X_1, X_2, \ldots, X_n$ مسحوبة من مجتمع معرف بدالة الكثافة الاحتمالية $f(x)$ $a < x < b$, هو :

$$g(y_1, y_2, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i) ; a < y_1 < y_2 < \ldots < y_n < b$$

وسوف نبرهن ذلك من خلال الفرض بان $n = 3$ والبرهان ذاته ينطبق على اية قيمة اخرى الى n .

البرهان : ان التوزيع المشترك الى X_1, X_2, X_3 هو

$$f(x_1, x_2, x_3) = f(x_1) . f(x_2) . f(x_3)$$

الان بفرض ان A تمثل مجموعة القيم x_1, x_2, x_3 . وحيث اننا بصدد ترتيب هذه القيم الثلاثة من اصغر قيمة في A الى اكبر قيمة فيها فذلك يعني اننا نتمكن من ترتيب هذه القيم بستة طرق مختلفة (اي $3!$) بحيث ان كل ترتيب منها يتمثل بمجموعة غير مرتبطة بمجموعة ترتيب آخر . هذه المجموعات الستة هي : . .

$$A_1 = \{(x_1, x_2, x_3) : a < x_1 < x_2 < x_3 < b\}$$
$$A_2 = \{(x_1, x_2, x_3) : a < x_2 < x_1 < x_3 < b\}$$
$$A_3 = \{(x_1, x_2, x_3) : a < x_1 < x_3 < x_2 < b\}$$
$$A_4 = \{(x_1, x_2, x_3) : a < x_2 < x_3 < x_1 < b\}$$
$$A_5 = \{(x_1, x_2, x_3) : a < x_3 < x_1 < x_2 < b\}$$
$$A_6 = \{(x_1, x_2, x_3) : a < x_3 < x_2 < x_1 < b\}$$

ولنفرض انه تم ترتيب القيم x_1 , x_2 , x_3 وفق ترتيب تصاعدي ، عندئذ فان القيم المرتبة المقابلة لها ستكون $a < y_1 < y_2 < y_3 < b$ وهذا يعني ان

$$y_1 = \min . (x_1 , x_2 , x_3)$$
$$y_2 = \text{mid} \ (x_1 , x_2 , x_3)$$
$$y_3 = \max . (x_1 , x_2 , x_3)$$

ان الدوال الثلاث y_1 , y_2 , y_3 في هذه الحالة سوف تعرف تحويل يؤدي الى تقابل عنصر مع عنصر مابين عناصر المجموعات A_i وعناصر المجموعة $B = \{ (y_1 , y_2 , y_3) : a < y_1 < y_2 < y_3 < b \}$ وهذا يعني ان :

لعناصر المجموعة A_1 سوف يكون

$$x_1 = y_1 \qquad , x_2 = y_2 \qquad , x_3 = y_3$$

لعناصر المجموعة A_2 سوف يكون

$$x_2 = y_1 \qquad , x_1 = y_2 \qquad , x_3 = y_3$$

لعناصر المجموعة A_3 سوف يكون

$$x_1 = y_1 \qquad , x_3 = y_2 \qquad , x_2 = y_3$$

لعناصر المجموعة A_4 سوف يكون

$$x_2 = y_1 \qquad , x_3 = y_2 \qquad , x_1 = y_3$$

لعناصر المجموعة A_5 سوف يكون

$$x_3 = y_1 \qquad , x_1 = y_2 \qquad , x_2 = y_3$$

ولعناصر المجموعه A_6 سوف يكون

$$x_3 = y_1 \qquad , x_2 = y_2 \qquad , x_1 = y_3$$

وهذا يعني انه في كل حالة هنالك تحويل من X الى Y مما يتطلب الامر تحديد معامل التحويل (جاكوبيان) الذي يمثل بمحدد مصفوفة ذات مرتبة 3×3 عناصرها تمثل مشتقات جزئية الى x_i نسبة الى y_i . وفي ضوء التقسيم اعلاه سوف نحصل على ست معاملات تحويل كل منها يتمثل بالآتي :

$$J_i = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} & \frac{\partial x_3}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} & \frac{\partial x_3}{\partial y_2} \\ \frac{\partial x_1}{\partial y_3} & \frac{\partial x_2}{\partial y_3} & \frac{\partial x_3}{\partial y_3} \end{vmatrix} \quad , i = 1, 2, \ldots, 6$$

فللمجموعة A_1 فان

$$J_1 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \rightarrow |J_1| = 1$$

وللمجموعة A_2 فان

$$J_2 = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -1 \rightarrow |J_2| = 1$$

ووفق نفس الاسلوب يمكن حساب بقية معاملات التحويل مع ملاحظة ان قيمة J_i هي $(+1)$ او (-1) الا انه في النتيجة $|J_i| = 1$. عليه فان

$$g(y_1, y_2, y_3) = \sum_{A_i \to B} f(x_1, x_2, x_3) \cdot |J_i|$$

$$= \sum_{A_i \to B} f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot |J_i|$$

$$= f(y_1).f(y_2).f(y_3) + f(y_2).f(y_1).f(y_3) + \ldots + f(y_3).f(y_2).f(y_1)$$

$$= 6f(y_1).f(y_2).f(y_3) = 3! \prod_{i=1}^{3} f(y_i) \, , a < y_1 < y_2 < y_3 < b$$

عليه وبشكل عام فان

$$g(y_1, y_2, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i) \, ; a < y_1 < y_2 < \ldots < y_n < b$$

١٠ ـ ٢ : توزيعات الاحصاءات المرتبة :

نستعرض في هذه الفقرة اسلوب اشتقاق الدوال الاحتمالية للاحصاءات المرتبة . وسوف نركز الاهتمام على اشتقاق دالة التوزيع العام للقيمة المرتبة y_k ، $k = 1, 2, \ldots, n$ ومنها نشتق دالة القيمة الصغرى y_1 ودالة القيمة العظمى y_n.

١٠ ـ ٢ ـ ١ : التوزيع العام لقيمة المرتبة y_k

لاحظنا في الفقرة (١٠ ـ ٢) ان التوزيع المشترك للاحصاءات المرتبة $y_1 < y_2 < \ldots < y_k$ هو $g(y_1, y_2, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i)$. وهذا يعني انه يمكن ايجاد التوزيع الحدي للمتغير Y_k من خلال اجراء التكامل لكافة المتغيرات الاخرى عدا Y_k ان التوزيع الحدي الى Y_k معطى بالصيغة التالية :

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!} [F(y_k)]^{k-1} . [1 - F(y_k)]^{n-k} . f(y_k)$$

البرهان :

$$g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2} . \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} n! \prod_{i=1}^{n} f(y_i) . dy_n \ldots dy_{k+1} . dy_1 \ldots dy_{k-1}$$

حيث ان : $a < y_1 < y_2 < \ldots < y_{k-1} < y_k < y_{k+1} < \ldots < y_n < b$

$$\therefore g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2} . \int_{y_k}^{b} \cdots \int_{y_{n-2}}^{b} n! \left[\int_{y_{n-1}}^{b} f(y_n) dy_n \right] . \prod_{i=1}^{n-1} f(y_i) .$$

$$dy_{n-1} \ldots dy_{k+1} . dy_1 \ldots dy_{k-1}$$

لكن

$$\int_{y_{n-1}}^{b} f(y_n)\,dy_n = \int_{y_{n-1}}^{b} dF(y_n) = F(y_n)\Big]_{y_{n-1}}^{b}$$

$$= F(b) - F(y_{n-1}) = 1 - F(y_{n-1})$$

$$g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2} . \int_{y_k}^{b} \cdots \int_{y_{n-2}}^{b} n!\,[1 - F(y_{n-1})] . \prod_{i=1}^{n-1} f(y_i) .$$

$$dy_{n-1} \dots dy_{k+1} . dy_1 . \dots dy_{k-1}$$

$$= \int_a^{y_k} \cdots \int_a^{y_2} . \int_{y_k}^{b} \cdots \int_{y_{n-3}}^{b} n! \left\{ \int_{y_{n-2}}^{b} [1 - F(y_{n-1})] f(y_{n-1})\,dy_{n-1} \right\}$$

$$\prod_{i=1}^{n-2} f(y_i) . dy_{n-2} \dots dy_{k+1} . dy_1 \dots dy_{k-1}$$

لكن

$$\int_{y_{n-2}}^{b} [1 - F(y_{n-1})] f(y_{n-1})\,dy_{n-1} = \int_{y_{n-2}}^{b} [1 - F(y_{n-1})] . dF(y_{n-1})$$

$$= -\frac{1}{2}[1 - F(y_{n-1})]^2 \Big]_{y_{n-2}}^{b} = \frac{1}{2}[1 - F(y_{n-2})]^2$$

$$\therefore\ g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2} . \int_{y_k}^{b} \cdots \int_{y_{n-3}}^{b} \frac{n!}{2!}[1 - F(y_{n-2})]^2 . \prod_{i=1}^{n-2} f(y_i) .$$

$$dy_{n-2} \dots dy_{k+1} . dy_1 . \dots dy_{k-1}$$

ولو استمرت عمليات التكامل بنفس الاجراء السابق نزولاً لغاية المتغير y_{k+1} فان

$$g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2} \frac{n!}{(n-k)!} [1 - F(y_k)]^{n-k} . \prod_{i=1}^{k} f(y_i).dy_1 \ldots dy_{k-1} ,$$

$$k < n$$

$$= \frac{n!}{(n-k)!} [1 - F(y_k)]^{n-k} . \int_a^{y_k} \cdots \int_a^{y_2} \prod_{i=1}^{k} f(y_i) dy_1 . dy_2 \ldots dy_{k-1}$$

$$= \frac{n!}{(n-k)!} [1 - F(y_k)]^{n-k} . \int_a^{y_k} \cdots \int_a^{y_3} \left[\int_a^{y_2} f(y_1) dy_1 \right] . \prod_{i=2}^{k} f(y_i) . dy_2 \ldots dy_{k-1}$$

لكن

$$\int_a^{y_2} f(y_1) dy_1 = \int_a^{y_2} dF(y_1) = [F(y_1)]_a^{y_2}$$

$$= F(y_2) - F(a) = F(y_2)$$

فاذن

$$g_k(y_k) = \frac{n!}{(n-k)!} [1 - F(y_k)]^{n-k} . \int_a^{y_k} \cdots \int_a^{y_3} . F(y_2) \prod_{i=2}^{k} f(y_i) . dy_2 \ldots dy_{k-1}$$

$$= \frac{n!}{(n-k)!} [1 - F(y_k)]^{n-k} . \int_a^{y_k} \cdots \int_a^{y_4} \left[\int_a^{y_3} F(y_2) . f(y_2) dy_2 \right] \prod_{i=3}^{k} f(y_i) . dy_3 \ldots dy_{k-1}$$

لكن

$$\int_a^{y_3} F(y_2) . f(y_2) dy_2 = \int_a^{y_3} F(y_2) . dF(y_2)$$

$$= \frac{[F(y_2)]^2}{2}\Big]_a^{y_3} = \frac{1}{2}[F(y_3)]^2$$

$$\therefore\ g_k(y_k) = \frac{n!}{(n-k)!}[1-F(y_k)]^{n-k}.\int_a^{y_k}\cdots\int_a^{y_4}\frac{1}{2!}[F(y_3)]^2.$$

$$\prod_{i=3}^{k} f(y_i).dy_3 \ldots dy_{k-1}$$

ولو استمرت عمليات التكامل بنفس الاجراء السابق صعوداً لغاية المتغير y_{k-1} فان

$$g_k(y_k) = \frac{n!}{(n-k)!}[1-F(y_k)]^{n-k}.\frac{1}{(k-1)!}[F(y_k)]^{k-1}.f(y_k)$$

فاذن

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!}.[F(y_k)]^{k-1}.[1-F(y_k)]^{n-k}.f(y_k)$$

$$a < y_k < b$$

$$\therefore\ g_k(y_k) = n\,C_{k-1}^{n-1}.[F(y_k)]^{k-1}.[1-F(y_k)]^{n-k}.f(y_k)\ \ldots(*)$$

ويلاحظ من (*) ان معامل الدالة $f(y_k)$ يمثل دالة توزيع ذي الحدين بالمعلمتين $(P = F(y_k)$ و $n-1)$ مضروبة بالعدد n .

١٠ ــ ٣ ــ ٢ : توزيع القيمة المرتبة الصغرى y_1

في حالة اختيارنا $k = 1$ فان الصيغة (*) تختزل الى الدالة التالية :

$$g_1(y_1) = n[1-F(y_1)]^{n-1}.f(y_1)\,;\,a < y_1 < b$$

والدالة g_1 تمثل الان التوزيع الاحتمالي للقيمة المرتبة الصغرى مع ملاحظة ان دوال الاحصاءات المرتبة تبقى تمتلك نفس خصائص دوال الكثافة الاحتمالية من حيث كونها دوال وحيدة القيمة ، موجبة، التكامل حول فضاء تلك القيمة المرتبة يجب ان يكون مساويا للواحد فمثلاً للدالة $g_1(y_1)$ فان :

$$\int_a^b g_1(y_1)\,dy_1 = n\int_a^b [1 - F(y_1)]^{n-1}.f(y_1)\,dy_1$$

$$= n\int_a^b [1 - F(y_1)]^{n-1}.dF(y_1)$$

$$= -n.\frac{[1 - F(y_1)]^n}{n}\Big]_a^b$$

$$= -[1 - F(y_1)]^n\Big]_a^b$$

$$= -\{[1 - F(b)]^n - [1 - F(a)]^n\} = -(-1) = 1$$

حيث ان $F(a) = 0$, $F(b) = 1$.

١٠ ـ ٣ ـ ٣ : توزيع القيمة المرتبة العظمى y_n .

في حالة اختيارنا $k = n$ فان الصيغة (*) تختزل الى الدالة التالية :

$$g_n(y_n) = n[F(y_n)]^{n-1}.f(y_n)\,;\, a < y_n < b$$

ونترك للقاريء البيان ان $g_n(y_n)$ هي دالة كثافة احتمالية .

وعلى ضوء التوزيع الاحتمالي للقيمة المرتبة y_k يتم حساب عزوم التوزيع المعرف بالدالة $g_k(y_k)$ وغيرها من الامور ذات العلاقة بالتوزيع (كالدالة التوزيعية ، حساب احتمال معين ، المنوال ، ... الخ) . فمثلا لتوزيع القيمة العظمى فان متوسط التوزيع سيعرف بالشكل التالي :

$$EY_n = \int_a^b y_n \cdot g_n(y_n)\, dy_n$$

وان تباين هذا التوزيع سيكون معرفاً بالشكل التالي :

$$V(Y_n) = EY_n^2 - (EY_n)^2$$

١٠ ـ ٣ ـ ٤ : التوزيع المشترك لاي قيمتين مرتبتين .

ليكن $a < y_1 < y_2 < \ldots < y_i < \ldots < y_j < \ldots < y_n < b$ يمثل ترتيب لقياسات عينة عشوائية قوامها n مسحوبة من مجتمع معرف بالدالة $f(x)$ ، $a < x < b$. وافرض اننا نرغب في ايجاد التوزيع المشترك للقيمتين المرتبتين $y_i , y_j , i < j$. عندئذ فان هذا الامر ممكن من خلال اجراء عمليات التكامل للتوزيع المشترك $g(y_1, y_2, \ldots, y_n)$ على المتغيرات كافة باستثناء المتغيرين Y_i و Y_j وكما هو مبين بالاتي :

$$g_{ij}(y_i,y_j) = \int_a^{y_2} \cdots \int_a^{y_i} \cdot \int_a^{y_{i+2}} \cdots \int_{y_{j-2}}^{b} \cdot \int_{y_j}^{b} \cdots \int_{y_{n-1}}^{b} n!. \prod_{h=1}^{n} f(y_h) dy_n \ldots$$

$$dy_{j+1} \cdot dy_{j-1} \ldots dy_{i+1} \cdot dy_{i-1} \ldots dy_1$$

وباجراء عمليات التكامل بدءاً بالمتغير Y_n نزولاً لغاية المتغير Y_{j+1} وبدءاً بالمتغير Y_{j-1} نزولاً لغاية المتغير Y_{i+1} وبدءاً بالمتغير Y_{i-1} نزولاً لغاية المتغير Y_1 . ووفق الاسلوب الذي اتبعناه في الفقرة (١٠ ـ ٣ ـ ١) فاننا سوف نحصل على :

$$g_{ij}(y_i,y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(y_i)]^{i-1} . [F(y_j) -$$

$$F(y_i)]^{j-i-1} . [1 - F(y_j)]^{n-j} . f(y_i) . f(y_j)$$

وان $a < y_i < y_j < b$. واذا فرضنا ان

$$Z_1 = i - 1 , Z_2 = j - i - 1 , Z_3 = n - j \rightarrow Z_1 + Z_2 + Z_3 = n - 2$$

وان

$$F(y_i) = P_1 \quad , F(y_j) - F(y_i) = P_2 , 1 - F(y_j) = P_3$$

فان

$$P_1 + P_2 + P_3 = 1$$

عليه فان

$$g_{ij}(y_i, y_j) = n(n-1) \frac{(n-2)!}{Z_1 ! Z_2 ! Z_3 !} . P_1^{Z_1} . P_2^{Z_2} . P_3^{Z_3} . f(y_i) . f(y_j)$$

ونلاحظ من الصيغة الاخيرة ان معامل $f(y_i) . f(y_j)$ يمثل توزيع متعدد الحدود، بثلاث متغيرات مضروب بالعدد $n(n-1)$.

وفي ضوء التوزيع المشترك لاي قيمتين مرتبتين يمكن الحصول على عزوم التوزيع المشتركة واية امور اخرى ذات علاقة بهذا التوزيع (كالدالة التوزيعية المشتركة ، العزوم المشتركة ، حساب احتمال مشترك معين ... الخ) فمثلاً لو تطلب الامر حساب EY_iY_j فان ذلك يتم وفق الاتي :

$$EY_iY_j = \int_a^b \int_a^b y_i y_j . g_{ij}(y_i, y_j) \, dy_i dy_j$$

كذلك فان

$$\sigma_{ij} = Cov(Y_i, Y_j) = EY_iY_j - EY_i . EY_j$$

وهذا يعني ان

معامل الارتباط بين القيمتين المرتبتين Y_j , Y_i هو $\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$

١٠ ـ ٣ ـ ٥ : مثال

افرض ان X_1, X_2, X_3, X_4, X_5 تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذا توزيع بيتا بالمعلمتين $\beta = 1 , \alpha = 2$.وافرض ان $Y_1 < Y_2 < Y_3 < Y_4 < Y_5$ يمثل الترتيب التصاعدي لقياسات هذه العينة . جد :

أ ـ توزيع القيمة العظمى ثم احسب توقع هذه القيمة

ب ـ توزيع القيمة الصغرى ثم احسب $P_r (Y_1 < 0.2)$.

جـ ـ توزيع القيمة الوسطى ثم جد تباين هذه القيمة .

د ـ التوزيع المشترك الى Y_4 , Y_2 ثم جد معامل الارتباط بينهما .

الحل : حيث ان $X \sim beta (2 , 1)$ فاذن :

$$f (x) = 2x , F (x) = x^2 ; 0 < x < 1$$

وهذا يعني ان :

$$f (y_i) = 2y_i , F (y_i) = y_i^2 , i = 1 , 2 , \ldots , 5 , n = 5$$

عليه فان :

أ ـ توزيع القيمة العظمى هو :

$$g_5 (y_5) = 5 [y_5^2]^4 . 2y_5 = 10y_5^9 , 0 < y_5 < 1$$

$$\therefore EY_5 = 10 \int_0^1 y_5^{10} dy_5 = \frac{10}{11} [y_5^{11}]_0^1 = \frac{10}{11}$$

ب ـ توزيع القيمة الصغرى هو :

$$g_1 (y_1) = 5 [1 - y_1^2]^4 . 2y_1 = 10y_1 (1 - y_1^2)^4 ; 0 < y_1 < 1$$

$$\therefore P_r (Y_1 \leq 0.2) = 10 \int_0^{0.2} y_1 (1 - y_1^2)^4 dy_1$$

$$= - (1 - y_1^2)^5]_0^{0.2} = 0.185$$

جـ ـ توزيع القيمة الوسطى هو :

$$g_3 (y_3) = \frac{5!}{2! . 2!} . [y_3^2]^2 . [1 - y_3^2]^2 . 2y_3$$

$$= 60y_3^5 (1 - y_3^2)^2 \ ; 0 < y_3 < 1$$

ويترك للقاريء حساب $V(Y_3)$.

د ـ التوزيع المشترك الى Y_4, Y_2 هو :

$$g_{24}(y_2, y_4) = \frac{5!}{1!.1!.1!}[y_2^2][y_4^2 - y_2^2].[1 - y_4^2].(2y_2)(2y_4)$$

$$= 480y_2^3y_4[y_4^2 - y_2^2][1 - y_4^2]; 0 < y_2 < y_4 < 1$$

ونترك للقاريء حساب ρ_{24}.

١٠ ـ ٤ : توزيعات دوال الاحصاءات المرتبة

في هذه الفقرةسوف نستعرض توزيعات دوال الاحصاءات المرتبة (اي توزيعات تلك الدوال التي تعتمد على القيم المرتبة) وفيما يخص الامر بـ الوسيط ، المدى ، منتصف المدى ، المدى القياسي .

١٠ ـ ٤ ـ ١ : تعاريف

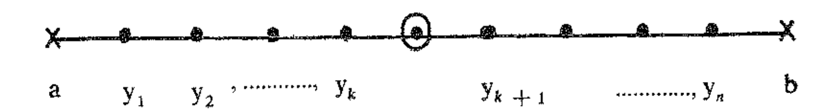
قبيل الدخول في موضوع توزيعات دوال الاحصاءات المرتبة لابد من اعطاء تعريف لكل مفهوم من المفاهيم الاربعة الواردة اعلاه . فاذا كان $a < Y_1 < Y_2 < \dots < Y_n < b$ يمثل الترتيب التصاعدي لقياسات عينة مسحوبة من مجتمع بدالة كثافة احتمالية $a < x < b, f(x)$ ، عندئذٍ ووفق هذه المعطيات يمكن تعريف هذه المفاهيم على النحو التالي :

١ ـ الوسيط Median

يعرف الوسيط بانه تلك القيمة من قيم العينة المرتبة تصاعدياً (او تنازلياً) التي تقسم مجموعة القيم المرتبة الى قسمين متساويين بحيث ان نصف القيم المرتبة تقع الى يمينها والنصف الاخر تقع الى يسارها . فاذا كان عدد القيم n عدد فردي عندئذٍ فان الوسيط يمثل القيم المرتبة ذات التسلسل $(n + 1)/2$. اما اذا كان عدد القيم n عدد زوجي فان الوسيط يمثل الوسط الحسابي للقيمتين المرتبتين ذات التسلسل $n/2, (n/2) + 1$. والشكل (١٠ ـ ١) يبين موقع الوسيط لكل حاله :

a y_1 y_2 ,............, $y_{(n/2)+1}$,............. y_n b

n عدد فردي



n عدد زوجي , $k = n/2$

الشكل (١٠ – ١) توضيح لموقع الوسيط

٢ ـ المدى Range

يعرف المدى بان الفرق ما بين اكبر قيمة في المجموعة المرتبة واصغر قيمة فيها . فاذا كان R يمثل المدى فان $R = Y_n - Y_1$.

٣ ـ منتصف المدى Mid-Range

يعرف منتصف المدى بانه متوسط المسافة بين Y_n , Y_1 لمجموعة قيم مرتبة . فاذا كان M يمثل منتصف المدى فان $M = (Y_1 + Y_n)/2$.

٤ ـ المدى القياسي Studentized Range

لتكن $X_1 , X_2 , \dots , X_n$ تمثل قياسات عينه عشوائية من $N(0,1)$ ، وان R يمثل المدى لهذه القياسات بعد ترتيبها تصاعدياً ليكن $V \sim \chi^2_{(m)}$. وبفرض ان V^2 مستقل عن قياسات العينة فان $S = R/(v/m)^{\frac{1}{2}}$ يسمى المدى القياسي .

١٠ ـ ٤ ـ ٢ : توزيع الوسيط Distribution of median

عندما تكون n عدداً فردياً فذلك يعني ان تسلسل الوسيط هو $K = (n + 1)/2$ وهذا يعني ان توزيع الوسيط لقياسات العينة يمثل الدالة (*) المنوه عنها في نهاية الفقرة (١٠ ـ ٣ ـ ١) بعد التعويض عن K بـ $(n + 1)/2$. والمثال الوارد في الفقرة (٩ ـ ٣ ـ ٥) الحالة (جـ) يمثل عملية ايجاد توزيع الوسيط للعينة المسحوبة من $f(x) = 2x$.

اما اذا كانت n عدداً زوجياً فذلك يتطلب اولاً حساب التوزيع المشترك للقيمتين المرتبتين ذات التسلسل $n/2 , (n/2) + 1$ اي التوزيع المشترك الى $Y_k , Y_{k+1} , K = n/2$. هذا التوزيع ناتج من تعويض $i = K , j = K + 1$ في التوزيع المشترك $g_{ij}(y_i , y_j)$ المنوه عنه في نهاية الفقرة (١٠ ـ ٣ ـ ٤).

وهذا يعني أن :

$$g(y_k \cdot y_{k+1}) = \frac{n!}{(K-1)!.(n-k-1)!} \cdot [F(y_k)]^{K-1}$$

$$[1 - F(y_{k+1})]^{n-k-1} . f(y_k) . f(y_{k+1})$$

وعندئذٍ وعلى اساس التحويل (حسب تعريف الوسيط) $Z = (y_k + y_{k+1})/2$ ، حيث Z تعني الوسيط للعينة ، يمكن استنتاج التوزيع الاحتمالي الى Z وكما هو مبين بالآتي :

$$Z = \frac{y_k + y_{k+1}}{2} \rightarrow a < Z < b - a$$

وبفرض أن

$$V = y_k \rightarrow y_k = v , a < v < b - Z$$

وان $y_{k+1} = 2Z - v$. الان معامل التحويل (جاكوبيان) هو .

$$J = \begin{vmatrix} \frac{\partial y_k}{\partial z} & \frac{\partial y_k}{\partial v} \\ \frac{\partial y_{k+1}}{\partial z} & \frac{\partial y_{k+1}}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 2 & -1 \end{vmatrix} = -2 \rightarrow |J| = 2$$

فاذن

$$g(z, v) = g(y_k, y_{k+1})\Big]_{\substack{y_k = v \\ y_{k+1} = 2z - v}} \cdot |J|$$

$$\therefore g(z, v) = \frac{2n!}{(k-1)!(n-k-1)!} [F(V)]^{k-1}$$

$$[1 - F(2z - v)]^{n-k-1} . f(v) . f(2z - v)$$

عليه فان

$$g(z) = \frac{2n!}{(k-1)!(n-k-1)!} \int_a^{b-z} [F(V)]^{k-1} .$$

$$[1 - F(2z - v)]^{n-k-1} f(v) . f(2z - v) \, dv$$

١٠ ــ ٤ ــ ٣ : توزيع المدى Distribution of Range:

لاحظنا ان المدى $R = Y_n - Y_1$ ، وبهدف تحديد التوزيع الاحتمالي الى R يتوجب اولاً ايجاد التوزيع المشترك بين Y_n, Y_1 . ان التوزيـع المشـترك $g_{1n}(y_1, y_n)$ يمكن الحصول عليه من خلال التعويض عن i. بـ i وعن j. بـ n في دالة التوزيع المشترك التي حصلنا عليها في الفقرة (١٠ ــ ٣ ــ ٤) . اي ان

$$g_{1n}(y_1, y_n) = n(n-1)[F(y_n) - F(y_1)]^{n-2} f(y_1) . f(y_n) ;$$

$$a < y_1 < y_n < b$$

وحيث ان $R = y_n - y_1$ فان $0 < R < b - a$. وبفرض ان $Z = y_1$ اي ان $y_1 = z$ فان $y_n = R + Z$ عليه فان معامل التحويل (جاكوبيان) J هو :

$$J = \begin{vmatrix} \frac{\partial y_1}{\partial R} & \frac{\partial y_1}{\partial Z} \\ \frac{\partial y_n}{\partial R} & \frac{\partial y_n}{\partial Z} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = -1 \rightarrow |J| = 1$$

فاذن

$$g(R, Z) = g(y_1, y_n)\Big]_{\substack{y_1 = Z \\ y_n = R+Z}} \cdot |J|$$

$$\therefore g(R, Z) = n(n-1)[F(R+Z) - F(Z)]^{n-2} . f(Z) f(R+Z)$$

عليه فان :

$$g(R) = n(n-1) \int_a^{b-R} [F(R+Z) - F(Z)]^{n-2} . f(Z) . f(R+Z) dz$$

الشكل (١٠ ــ ٢) ، توضيح للقيم الممكنة الى Z, R.

مثال : افرض ان X_1, X_2, X_3 تمثل عينة عشوائية من توزيع بيتا بالمعلمتين $\beta = 1, \alpha = 2$ ، وان $Y_1 < Y_2 < Y_3$ يمثل ترتيب قياسات هذه العينة . جد التوزيع الاحتمالي الى R ثم احسب متوسط التوزيع وتباينه .

الحل :

ان $f(x) = 2x \, ; 0 < x < 1 \qquad \therefore F(x) = x^2$

هذا يعني ان

$$f(y_i) = 2y_i \, , F(y_i) = y_i^2 \, ; 0 < y_1 < y_2 < y_3 < 1$$

$$\therefore f(Z) = 2Z \, , f(R + Z) = 2(R + Z)$$

$$F(Z) = Z^2 \, , F(R + Z) = (R + Z)^2$$

$$\therefore g(R) = 3(3 - 1) \int_0^{1-R} [(R + Z)^2 - Z^2] . (2Z) [2(R + Z)] dz$$

$$= 24 \int_0^{1-R} Z(R + Z)(R^2 + 2RZ) dz$$

وبفتح الاقواس واجراء عملية التكامل نحصل على

$$g(R) = 12 R (1 - R)^2 \, ; 0 < R < 1$$

ويتضح من هذا المثال ان $R \sim beta(2,3)$. فاذن : $\alpha = 2, \beta = 3$

$$ER = \frac{\alpha}{\alpha + \beta} = \frac{2}{5}$$

$$V(R) = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)} = \frac{1}{25}$$

١٠ ـ ٤ ـ ٤ : توزيع منتصف المدى Distribution of mid-range

لاحظنا من تعريف منتصف المدى M ان $M = (Y_1 + Y_n)/2$. ولغرض ايجاد التوزيع الاحتمالي الى M يتوجب اولاً ايجاد التوزيع المشترك الى Y_1, Y_n هذا التوزيع سبق وان وجدناه في الفقرة السابقة وكان :

$$g_{1n}(y_1, y_n) = n(n-1)[F(y_n) - F(y_1)]^{n-2} \cdot f(y_1) \cdot f(y_n)$$

الآن

$$M = \frac{y_1 + y_n}{2} \rightarrow a < M < b - a$$

وبفرض ان $V = y_1$ فان $y_1 = V$. عليه فان $y_n = 2M - V$ وان $a < V < b - M$ وان معامل التحويل J هو :

$$J = \begin{vmatrix} \frac{\partial y_1}{\partial M} & \frac{\partial y_1}{\partial V} \\ \frac{\partial y_n}{\partial M} & \frac{\partial y_n}{\partial V} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 2 & -1 \end{vmatrix} = -2 \rightarrow |J| = 2$$

عليه فان

$$g(M, V) = g(y_1, y_n)\Big]_{\substack{y_1 = V \\ y_n = 2M - V}} \cdot |J|$$

$$\therefore g(M, V) = 2n(n-1)[F(2M - V) - F(V)]^{n-2} \cdot f(V) \cdot f(2M - V)$$

$$\therefore g(M) = 2n(n-1) \int_a^{b-M} [F(2M - V) - F(V)]^{n-2} \cdot f(V) \cdot f(2M - V)\, dV$$

مثال : لمعطيات المثال الوارد في الفقرة (١٠ ـ ٤ ـ ٣) جد توزيع منتصف المدى

الحل :

$$f(V) = 2V, \; f(2M - V) = 2(2M - V)$$

$$F(V) = V^2, \; F(2M - V) = (2M - V)^2 \,;\, 0 < V < 1 - M, \quad 0 < M < 1$$

$$\therefore g(M) = 12 \int_0^{1-M} [(2M-V)^2 - V^2].$$

$$(2V)[2(2M-V)]\,dV$$

$$= 48 \int_0^{1-M} V(2M-V)[(2M-V)^2 - V^2]\,dV$$

وبفتح الاقواس واجراء عملية التكامل نحصل على :

$$g(M) = 48\,M[(3M-1)(1-M)]^2 \;; 0 < M < 1$$

١٠ ـ ٤ ـ ٥ : توزيع المدى القياسي
Distribution of studentized range

من تعريف المدى القياسي لاحظنا ان :

$$S = \frac{R}{\sqrt{\frac{V}{m}}} \;,\; V \sim \chi^2_{(m)} \;,\; 0 < V < \infty$$

بشكل عام فان $0 < R < \infty$ لاي توزيع احتمالي معرفـة قيمة فـي الفتـرة $[-\infty,\infty]$ وهذا يعني ان $0 < S < \infty$. وبغية حساب التوزيع الاحتمالي للمتغير S يتوجب اولاً ايجاد التوزيع المشترك الى R و V . اي $g(R,V)$. وحيث ان V مستقل عن قياسات العينة X_i فذلك يعني ان V مستقل عن R. لذا فان $g(R,V) = g_1(R)\,g_2(V)$ وان $g_1(R)$ معرفة في الفقرة (١٠ ـ ٤ ـ ٣). الان بفرض ان $u = V/m$ فان $S = R\sqrt{u}$. عليه فان $R = S\sqrt{u}$ وان $v = mu$. وحيث ان $0 < V < \infty$ فان $0 < u < \infty$ ان معامل التحويل J هو :

$$J = \begin{vmatrix} \frac{\partial R}{\partial S} & \frac{\partial R}{\partial u} \\ \frac{\partial V}{\partial S} & \frac{\partial V}{\partial u} \end{vmatrix} = \begin{vmatrix} \sqrt{u} & \frac{S}{2\sqrt{u}} \\ 0 & m \end{vmatrix} = m\sqrt{u}$$

$$\therefore |J| = m\sqrt{u}$$

$$\therefore f(S,u) = g(R,V)\Big]_{\substack{R = S\sqrt{u} \\ V = mu}} \cdot |J|$$

$$\therefore f(S,u) = g_1(S\sqrt{u}) \cdot g_2(mu) \cdot m\sqrt{u}$$

$$\therefore f(S) = m\int_0^\infty g_1(S\sqrt{u}) \cdot g_2(mu) \cdot \sqrt{u}\, du, S > 0$$

ان الدالة $f(S)$ تسمى دالة التوزيع الاحتمالي للمدى القياسي . وبشكل خاص اذ كان $X \sim N(0,1)$ ان $\phi(x)$ تمثل الدالة الاحتمالية الى X وان $\Phi(x)$ تمثل الدالة التوزيعية لهذا المتغير فان :

$$g(R) = n(n-1)\int_{-\infty}^{\infty} [\Phi(R+Z) - \Phi(Z)]^{n-2}\phi(Z) \cdot \phi(Z) \cdot \phi(R+Z)\, dZ$$

ومنها (ووفق نفس الاسلوب المتبع اعلاه) نصل الى

$$f(S) = m\int_0^\infty g_1(S\sqrt{u}) \cdot g_2(mu) \cdot \sqrt{u}\, du, S > 0$$

ان الدالة $f(S)$ بصيغتها المعتمدة على $N(0,1)$ هي الشائعة الاستخدام في بعض تطبيقات النظرية الاحصائية وخصوصاً في موضوعي اختبار الفرضيات وبناء حدود الثقة . وفي هذه الحالة يقال ان المتغير العشوائي S يتوزع كتوزيع المدى القياسي بدرجتي حرية n و m بدالة كثافة احتمالية $f(S)$. وقد تم بناء

جداول خاصة (لاحظ الجدول ٩ ملحق ب) بهذا التوزيع تبين قيم S النظرية عند مستوى معنوية α وبدرجتي حرية n ، m . فاذا كانت $S_{n,m}(\alpha)$ تمثل قيمة المدى القياسي عند درجتي حرية m,n ومستوى معنوية α فان هذه القيمة تحقق المعادلة التكاملية التالية :

$$\int_{S_{n,m}(\alpha)}^{\infty} f(S)\,dS = \alpha = P_r(S > S_{n,m}(\alpha))$$

اوان

$$P_r(S \leq S_{n,m}(\alpha)) = 1 - \alpha$$

فمثلاً قيمة $S_{10,5}$ التي تعطي احتمال متراكم قدره 0.95 هي 6.80 اي ان

$$P_r(S \leq 6{\cdot}80) = 0{\cdot}95$$

او ان

$$P_r(S > 6{\cdot}80) = 0{\cdot}05$$

ويلاحظ من جداول هذا التوزيع ان $n \geq 2$ بسبب ان عملية حساب المدى R تتطلب وجود عينة عشوائية عدد قياساتها لا يقل عن قياستين .

مثال : افرض ان X_1, X_2 تمثل عينة مسحوبة من $N(0,1)$ وان $Y_1 < Y_2$ يمثل الترتيب التصاعدي لهاتين القياستين . ليكن $V \sim \chi^2_{(1)}$. جد التوزيع الاحتمالي للمدى القياسي .

الحل : واضح من معطيات المثال ان :

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad , \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2}\, dw$$

وان

$$g(V) = \frac{1}{\sqrt{2\pi}} \cdot V^{-\frac{1}{2}} e^{-\frac{v}{2}} \quad , V \sim \chi^2_{(1)}, V > 0$$

فاذن

$$\phi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} , \phi(R+Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(R+Z)^2}$$

$$\therefore g(R) = 2\int_{-\infty}^{\infty} \phi(Z) . \phi(R+Z)\, dz$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2 - \frac{1}{2}(R+Z)^2} dz$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2 - \frac{1}{2}R^2 - RZ - \frac{1}{2}z^2} dz$$

$$= \frac{1}{\pi} e^{-\frac{1}{2}R^2} \int_{-\infty}^{\infty} e^{-(Z^2 + RZ)} dz$$

وباكمال المربع للكمية داخل القوسين نحصل على :

$$g(R) = \frac{1}{\pi} e^{-\frac{1}{2}R^2} . e^{\frac{1}{4}R^2} \int_{-\infty}^{\infty} e^{-(Z+\frac{1}{2}R)^2} dz$$

$$= \frac{1}{\pi} e^{-\frac{1}{4}R^2} . \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{Z+R/2}{1/\sqrt{2}}\right)^2} dz$$

واذا فرضنا التحويل التالي :

$y = \sqrt{2}\left(Z + \frac{R}{2}\right)$ فان $dy = \sqrt{2}\, dz$ أو أن $dz = \frac{1}{\sqrt{2}} dy$ وعليه فان

$$g(R) = \frac{1}{\sqrt{2}\,\pi} e^{-\frac{1}{4}R^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy$$

والتكامل في الصيغة الاخيرة مساو الى $\sqrt{2\pi}$. فاذن

$$g(R) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{4}R^2} \cdot \sqrt{2\pi}$$

$$= \frac{1}{\sqrt{\pi}} e^{-\frac{1}{4}R^2}, R > 0$$

والدالة الاخيرة تمثل التوزيع الاحتمالي للمدى في هذه العينة . عليه فان

$$f(s) = m \int_0^\infty g_1(s\sqrt{u}) \cdot g_2(mu) \cdot \sqrt{u}\, du$$

حيث ان :

$$g_1(s\sqrt{u}) = g(R)\Big]_{R = S\sqrt{u}} = \frac{1}{\sqrt{\pi}} e^{-\frac{1}{4}S^2 \cdot u}$$

وان

$$g_2(mu) = g(v)\Big]_{v = mu} = \frac{1}{\sqrt{2\pi}} \cdot u^{-\frac{1}{2}} \cdot e^{-\frac{u}{2}} \; ; m = 1$$

$$\therefore f(s) = \int_0^\infty \frac{1}{\sqrt{\pi}} e^{-\frac{1}{4}S^2 \cdot u} \cdot \frac{1}{\sqrt{2\pi}} u^{-\frac{1}{2}} \cdot e^{-\frac{u}{2}} \cdot \sqrt{u}\, du$$

$$= \frac{1}{\sqrt{2}\,\pi} \int_0^\infty e^{-\frac{1}{4}S^2 \cdot u - \frac{1}{2}u} du$$

$$= \frac{1}{\sqrt{2}\,\pi} \int_0^\infty e^{-\lambda u} du, \lambda = \frac{1}{2} + \frac{1}{4}S^2$$

$$= \frac{-1}{\lambda\sqrt{2}\,\pi} \left[e^{-\lambda u} \right]_0^\infty = \frac{1}{\lambda\sqrt{2}\,\pi}$$

$$\therefore\ f(S) = \frac{1}{\sqrt{2}\,\pi\left(\frac{1}{2} + \frac{1}{4}S^2\right)}$$

$$\therefore\ f(S) = \frac{1}{\frac{\pi}{\sqrt{2}}\left(1 + \frac{1}{2}S^2\right)} \quad , S > 0$$

ان الدالة $f(S)$ تسمى دالة التوزيع الاحتمالي للمدى القياسي S لهذه العينة . ويلاحظ ان :

$$\int_0^\infty f(S)\,ds = \frac{\sqrt{2}}{\pi}\int_0^\infty \frac{ds}{\left(1 + \frac{1}{2}S^2\right)}$$

وبفرض ان $y = \frac{S}{\sqrt{2}}$ فان $S = \sqrt{2}\,y$, $ds = \sqrt{2}\,dy$, $y > 0$. فاذن

$$\int_0^\infty f(S)\,ds = \frac{\sqrt{2}}{\pi}\int_0^\infty \frac{\sqrt{2}\,dy}{1 + y^2} = \frac{2}{\pi}\int_0^\infty \frac{dy}{1 + y^2}$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{dy}{1 + y^2} = \int_{-\infty}^{\infty} \frac{dy}{\pi(1 + y^2)} = 1$$

لاحظ ان التكامل الاخير يمثل تكامل دالة توزيع كوشي المعياري حول فضاء المتغير Y وكأن $Y \sim \text{cauchy}(0, 1)$.

تمارين

١٠ ــ ١ : افرض ان X_1, X_2, X_3, X_4 تمثل قياسات عينة عشوائية مسحوبة من مجتمع بدالة كثافة احتمالية $f(x) = 3x^2 ; 0 < x < 1$. جد مايلي :

أ ــ دالة التوزيع الاحتمالي لاصغر قيمة من قيم هذه العينة ثم جد متوسطها .

ب ــ دالة التوزيع الاحتمالي لاكبر قيمة من قيم هذه العينة ثم جد تباينها .

جـ ــ دالة التوزيع الاحتمالي المشترك بين Y_3 , Y_2 ثم جد ρ_{23}.

د ــ دالة التوزيع الاحتمالي لمدى هذه العينة ثم جد تباين المدى .

١٠ ــ ٢ : لتكن X_1, X_2, X_3, X_4, X_5 تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذا توزيع اسي بالمعلمة $\theta = 1$. جد مايلي :

أ ــ دالة التوزيع الاحتمالي لاصغر قيمة من قيم هذه العينة ثم جد تباينها.

ب ــ دالة التوزيع الاحتمالي لاكبر قيمة من قيم هذه العينة ثم جد متوسطها .

جـ ــ دالة التوزيع الاحتمالي لوسيط هذه العينة ثم جد القيمة المتوقعة للوسيط .

د ــ دالة التوزيع الاحتمالي لمدى هذه العينة ثم جد القيمة المتوقعة للمدى

١٠ ــ ٣ : لمعطيات السؤال الاول . جد التوزيع الاحتمالي لوسيط هذه العينة ثم جد متوسط وتباين الوسيط .

١٠ ــ ٤ : افرض ان X_1, X_2, X_3 تمثل قياسات عينة مسحوبة من مجتمع بدالة احتمالية $f(x) = 2x ; 0 < x < 1$ وافرض ان Z يمثل الوسيط لهذه العينة . جد $P_r[\min(X_i) > Z]$.

١٠ ــ ٥ : افرض ان R يمثل المدى لقياسات عينة عشوائية قوامها 4 مسحوبة من مجتمع ذا توزيع منتظم على الفترة $(0, 1)$. جد $P_r\left(R < \frac{1}{2}\right)$ ثم جد $V(R), ER$.

١٠ ــ ٦ : لتكن X_1, X_2, X_3 تمثل قياسات عينة مسحوبة من مجتمع ذي توزيع منتظم على الفترة $(0, 1)$. جد :

أ ــ التوزيع الاحتمالي لمنتصف المدى لهذه العينة ثم جد متوسطة وتباينه .

ب ــ التوزيع الاحتمالي للمدى القياسي لهذه العينة ثم جد متوسطه وتباينه .

١٠ ــ ٧ : لتكن $Y_1 < Y_2$ تمثل قياستين مسحوبتين من $N(0, \sigma^2)$ بين ان $EY_1 = -\sigma / \sqrt{\pi}$.

١٠ ــ ٨ : افرض ان $Y_1 < Y_2 < Y_3 < Y_4 < Y_5$ تمثل القيم المرتبة لعينة عشوائية مسحوبة من مجتمع ذا توزيع بيتا بالمعلمتين $\alpha = 3$, $\beta = 1$. وافرض ان $Z_1 = Y_2 / Y_4$, $Z_2 = Y_4$. بين ان Z_1 مستقل عن Z_2

١٠ ــ ٩ : لمعطيات السؤال (١٠ ــ ٥) . جد ما يلي :

أ ــ صيغة للعزم ذات المرتبة k حول نقطة الاصل الى R .

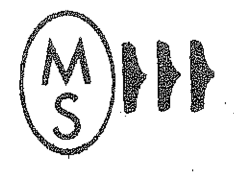
ب ــ جد المنوال في التوزيع الاحتمالي الى R .

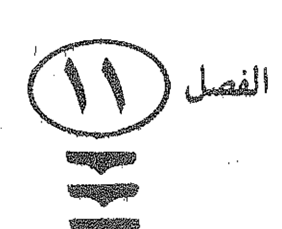
جـ ــ جد الدالة التوزيعية الى R ثم احسب قيمة $P_r(R > 0{\cdot}25)$.

د ــ جد الوسيط في التوزيع الاحتمالي الى R .

١٠ ــ ١٠ : لتوزيع القيمة العظمى $g_n(y_n)$. برهن ان الوسيط لهذا التوزيع هو العدد M الذي يحقق :

$$F(M) = \sqrt[n]{0{\cdot}5}$$





مقدمة في نظرية التقدير

الفصل الحادي عشر
مقدمة في نظرية التقدير

ان لنظرية التقدير (التخمين) Theory of estimation أهمية كبيرة في تطبيقات النظرية الاحصائية في الجوانب العملية . ومن وجهة نظر احصائية قسم علماء الاحصاء النظرية الاحصائية الى قسمين رئيسين هما الاحصاء الوصفي Descriptive Statistics الذي يهتم عادة باصول وقواعد جمع البيانات والمعلومات عن ظاهرة معينة أو مجموعة ظواهر والتركيز على عملية تصنيف وتبويب وعرض هذه البيانات وحساب بعض المؤشرات الاحصائية التي تخص معالم ذلك المجتمع الذي تم جمع البيانات من مفرداته سواء كان ذلك عن طريق التعداد الشامل لكافة مفردات المجتمع او عن طريق عينة مختارة منه . اما القسم الثاني فهو الاحصاء الاستدلالي Inferential Statistics الذي يهتم باصول وقواعد حساب افضل تقديرات لمعالم المجتمع من خلال نظرية التقدير وكذلك اختبار الفرضيات الخاصة بتلك المعالم (الذي سيأتي ذكره في الفصل القادم) .

وبشكل عام يمكن النظر الى نظرية التقدير على انها جزئين متكاملين الاول يهتم بالبحث عن افضل تقدير (مخمن) لمعلمة مجهولة تخص المجتمع . وهذا غالباً ما يسمى « التقدير بنقطة » " Point estimation " في حين ان الجزء الثاني يهتم بالبحث عن افضل فترة يمكن حصر قيمة المعلمة المجهولة خلالها . وهذا غالباً ما يسمى « التقدير بفترة » " Interval estimation " . وسوف يتم التطرق لكل جزء من هذين الجزئين بالشكل الذي لا يجعلنا ان نخرج عن نطاق وهدف هذا الكتاب كون ان القاريء سوف يدرس وبشكل مفصل هذه النظرية في مرحلة لاحقة وان هدفنا من عرض هذا الفصل هو تعريف القاريء بهذه النظرية دون الدخول في تفاصيلها الجزئية .

١١ ـ ١ : التقدير بنقطة Point Estimation

ان الهدف الاساس من هذه الفقرة هو استعراض لاهم الخواص التي يجب ان تتوفر في تقدير Estimator معين كي يمكننا اعتبار ذلك التقدير « أفضل تقدير » من بين جملة تقديرات اخرى متاحة فيما يخص معلمة (أو جملة معالم) مجهولة تخص مجتمع احصائي معين ، كأن تكون هذه المعلمة مثلاً متوسط عمر الطالب في جامعة الموصل (كمجتمع احصائي) ، متوسط انتاجية الدونم الواحد من الحنطة في المحافظات الشمالية من العراق (كمجتمع احصائي) وغيرها من الامثلة الاخرى . وبالتأكيد فان عملية حساب قيمة عددية لتلك المعلمة (او المعالم) من خلال الحصر الشامل لكافة مفردات المجتمع الاحصائي كاملة تقودنا الى القيمة الحقيقية لتلك المعلمة الامر الذي لا يقتضي التطرق لنظرية التقدير . الا ان مسألة حساب تلك القيمة من خلال الحصر الشامل يبدو امر شبه مستحيل لاعتبارات تتعلق بتكاليف الحصر الشامل من حيث المستلزمات المادية والبشرية والزمن اللازم لذلك . اضف الى ذلك انه في احوال كثيرة يتعذر حصر مفردات المجتمع الاحصائي كاملة كونه مجتمع غير محدود (كمجتمع الاسماك في بحيرة معينة) ، (مجتمع كريات الدم الحمراء في جسم الانسان) وغيرها من الامثلة كذلك في اختبارات السيطرة على نوعية الانتاج فان الامر يتطلب في احوال كثيرة تدمير الانتاج بهدف التعرف على نوعيته على سبيل المثال الاختبارات الخاصة بالعيارات النارية والقنابل . الاختبارات الخاصة بفحص نوعية الغزل القطني المنتج في مصنع للغزل والنسيج وغيرها من الامثلة .

وعلى ضوء ما تقدم فان الاسلوب الامثل البديل للحصر الشامل هو اسلوب العينات الذي يضمن توفيراً في الوقت والجهد والمال المبذولة في الحصر الشامل اضافة الى كونه الاسلوب الانسب في حالة المجتمعات غير المحدودة .

ان استخدام اسلوب العينات في حساب قيمة عددية لمعلمة (أو مجموعة معالم) تخص مجتمع الدراسة يعني الحصول على تقدير (او مجموعة تقديرات) غير مساو لقيمة المعلمة الحقيقية . فقد تكون قيمة التقدير اكبر من قيمة المعلمة أو اقل منها وذلك امر ناجم بسبب استخدامنا لجزء من المعلومات المتاحة في ذلك المجتمع .

وكما سبق ذكره في الفقرة (٨ ـ ١ ـ ٢) فان التقدير المستحصل عليه من العينة يعتبر هو الاخر بحكم متغير عشوائي بسبب اختلاف قيمته من عينة لاخرى ممكنة

السحب من ذلك المجتمع . فعلى فرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية ذات حجم n مسحوبة من مجتمع بدالة كتلة احتمالية $p(x, \theta)$ او دالة كثافة احتمالية $f(x, \theta)$ ، حيث ان θ تمثل معلمة تخص وتشخص ذلك المجتمع ، واننا نرغب وعلى أساس قياسات هذه العينة الوصول الى افضل تقدير ممكن الى θ مثل $\hat{\theta}$. فمثلاً اذا كانت هذه العينة مسحوبة من مجتمع ذي توزيع پواسون بالمعلمة λ فان الامر منصب على تقدير افضل قيمة عددية ، ممكنة الى $\theta = \lambda$. واذا كانت هذه العينة مسحوبة من $N(\mu, \sigma^2)$ فان الامر سيكون منصب على تقدير افضل قيمة عددية لكل من $\theta_1 = \mu$ و $\theta_2 = \sigma^2$. ان التقدير $\hat{\theta}$ سبق وان اسميناه في الفقرة (٨ ــ ١ ــ ٢) مؤشر الاحصائي Statistic وهو دالة بدلالة قياسات العينة لايعتمد على معلمة (او معالم) التوزيع ، أي ان $\hat{\theta} = g(x_1, x_2, \ldots, x_n)$ وان $\hat{\theta}$ يعتبر هو الاخر بحكم متغير عشوائي يسلك وفق دالة كتلة أو كثافة احتمالية معينة تعتمد بطبيعة الحال على حجم العينة n والتوزيع الاحتمالي للمتغير العشوائي X .

ان المبدأ العام لنظرية التقدير بنقطة هو التوصل الى افضل تقدير best estimator من بين جملة تقديرات اخرى ممكنة بحيث ان هذا التقدير يكون اقرب ما يمكن لقيمة المعلمة θ المطلوب تقديرها . وبغية اعتبار هذا التقدير كافضل تقدير ممكن فان ذلك يتطلب صفات معينة يتوجب توفرها في ذلك التقدير كي يسمح لنا ذلك تسمية $\hat{\theta}$ تقدير جيد . هذه الصفات هي الآتي :

١١ ــ ١ ــ ١ : الاتساق Consistency

افرض ان $\hat{\theta}_n$ يمثل تقدير للمعلمة θ على اساس عينة عشوائية قوامها n مسحوبة من مجتمع معرف بدالة كتلة احتمالية $p(x, \theta)$ أو دالة كثافة احتمالية $f(x, \theta)$. عندئذٍ يقال ان $\hat{\theta}_n$ هو تقدير متسق للمعلمة θ اذا كان $\hat{\theta}_n$ متقارب بالاحتمال من θ عندما $n \to \infty$ أي ان : $\hat{\theta}_n \xrightarrow{P} \theta$ عندما $n \to \infty$. وهذا يعني ان :

$$\lim_{n \to \infty} P_r(|\hat{\theta}_n - \theta| > \varepsilon) = 0 \quad \forall \varepsilon > 0$$

او ان

$$P_r(\lim_{n \to \infty} \hat{\theta}_n = \theta) = 1$$

ان هذه الصفة تعني فيما تعينه ان احتمال حدوث فرق مطلق بين $\theta, \hat{\theta}_n$ يتضاءل تدريجياً عندما يزداد حجم العينة n نحو عدد كبير جداً (نظرياً ∞).

مثال (١) : افرض ان $x_1, x_2, ..., x_n$ تمثل قياسات عينة عشوائية مسحوبة من توزيع پواسون بالمعلمة λ . ليكن $\bar{x} = \frac{1}{n} \Sigma x_i$ يمثل الوسط الحسابي لهذه العينة . بين ان $\bar{x}_n$ هو تقدير متسق الى λ .

الحل :

واضح هنا ان $\lambda = \theta$ وان $\hat{\theta}_n = \bar{x}_n$. كذلك وحيث ان القياسات $x_1, x_2, ..., x_n$ تمثل قياسات عينة عشوائية فذلك يعني انها بحكم متغيرات عشوائية مستقلة ذات نفس التوزيع أي ان :

$$X_i \sim \text{Poisson}(\lambda), i = 1, 2, ..., n$$

وان $V(\bar{x}_n) = \frac{\lambda}{n}$ عليه واستناداً لمتباينة تشيبيشيف فان :

$$P_r(|\bar{x}_n - \lambda| > \varepsilon) \leq \frac{\lambda}{n\varepsilon^2} \quad ; \varepsilon > 0$$

فاذن

$$\lim_{n \to \infty} P_r(|\bar{x}_n - \lambda| > \varepsilon) \leq \lim_{n \to \infty} \frac{\lambda}{n\varepsilon^2} = 0$$

وحيث ان الاحتمال لا يمكن ان يكون قيمة سالبة فاذن :

$$\lim_{n \to \infty} P_r(|\bar{x}_n - \lambda| > \varepsilon) = 0$$

وهذا يعني ان

$$P_r(\lim_{n \to \infty} \bar{x}_n = \lambda) = 1 .$$

اي ان $\bar{x}_n$ تقدير متسق الى λ .

مثال (٢) : افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$. ليكن $S_n^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$ يمثل تباين قياسات هذه العينة . بين ان S_n^2 هو تقدير متسق الى σ^2 .

الحل :

واضح هنا ان $\hat{\theta}_n = S_n^2 , \theta = \sigma^2$. واستناداً للفقرة (٩ ـ ١ ـ ١١) فان $V(S_n^2) = \frac{2\sigma^4}{n-1}$. عليه وحسب متباينة تشيبيشيف :

$$P_r(|S_n^2 - \sigma^2| > \varepsilon) \leq \frac{2\sigma^4}{(n-1)\varepsilon^2} , \varepsilon > 0$$

فاذن

$$\lim_{n \to \infty} P_r(|S_n^2 - \sigma^2| > \varepsilon) \leq \lim_{n \to \infty} \frac{2\sigma^4}{(n-1)\varepsilon^2} = 0$$

وهذا يعني ان

$$P_r(\lim_{n \to \infty} S_n^2 = \sigma^2) = 1$$

أي ان S_n^2 تقدير متسق الى σ^2 .

مما تقدم نلاحظ ان صفة الاتساق هي صفة غاية limit تعبر عن سلوك التقدير $\hat{\theta}_n$ عندما يزداد حجم العينة نحو عدد كبير (نظريا ∞). وهذا يعني ان هذه الصفة ليس لها أي معنى في حالة كون n محدودة كذلك وبافتراض ان $\hat{\theta}_n$ تقدير متسق الى θ فانه في ذات الوقت يوجد عدد غير منته من تقديرات مماثلة الى $\hat{\theta}_n$ تعتبر ايضاً تقديرات متسقة الى θ . فمثلاً التقديرات $\frac{n-a}{n-b} \hat{\theta}_n$ ، b, a ثوابت حقيقية ، هي ايضاً تقديرات متسقة الى θ .

مثال (٣) : افرض ان $\bar{x}_n$ يمثل تقدير متسق للمعلمة μ في $N(\mu, \sigma^2)$ وافرض ان $g_n = \frac{n-a}{n-b} \bar{x}_n$. بين ان g_n تقدير متسق الى μ

الحل :

استناداً لمتباينة تشيبيشسف فان :

$$P_r(|g_n - \mu| > \varepsilon) \leq \frac{V(g_n)}{\varepsilon^2}, \varepsilon > 0$$

وان

$$V(g_n) = \left(\frac{n-a}{n-b}\right)^2 . V(\bar{x}_n) = \left(\frac{n-a}{n-b}\right)^2 . \frac{\sigma^2}{n}$$

عليه فان

$$\lim_{n \to \infty} P_r(|g_n - \mu| > \varepsilon) \leq \lim_{n \to \infty} \left(\frac{n-a}{n-b}\right)^2 . \frac{\sigma^2}{n\varepsilon^2}$$

لكن

$$\lim_{n \to \infty} \left(\frac{n-a}{n-b}\right)^2 . \frac{\sigma^2}{n\varepsilon^2} = \lim_{n \to \infty} \left(\frac{1-a/n}{1-b/n}\right)^2 . \lim_{n \to \infty} \frac{\sigma^2}{n\varepsilon^2} = 0$$

فاذن

$$\lim_{n \to \infty} P_r(|g_n - \mu| > \varepsilon) = 0$$

أي ان

$$P_r(\lim_{n \to \infty} g_n = \mu) = 1$$

ويلاحظ من معطيات المثال (٣) ان التقدير المتسق لا يتصف بصفة الوحدانية حيث انه يوجد في ذات الوقت عدد غير منته من التقديرات المتسقة لنفس المعلمة كل منها يتحدد من خلال تخصيص قيمة لكل من b, a .

١١ ـ ١ ـ ٢ : عدم التحيز Unbiasedness

لاحظنا في الفقرة السابقة ان صفة الاتساق هي صفة غاية تعبر عن سلوك $\hat{\theta}_n$ عندما $n \to \infty$. وعلى العكس من ذلك فان صفة عدم التحيز تقترن بحالة كون n عدد محدود

يقال ان $\hat{\theta}_n$ تقدير غير متحيز الى θ اذا كان $E\hat{\theta}_n = \theta$. هذه الصفة تعني ان القيمة المتوقعة للتقدير $\hat{\theta}_n$ تساوي قيمة المعلمة θ . أو بمعنى آخر فان هذه الصفة تعني ان « متوسط »قيم التقديرات المستحصل عليها من كافة العينات الممكنة السحب من المجتمع فيما يخص المعلمة θ يجب ان يكون مساوياً الى θ . فاذا كان عدد تلك العينات هو r وان $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_r$ تمثل تقديرات هذه

العينات للمعلمة θ فان مفهوم صفة عدم التحيز هو ان $\frac{1}{r} \sum_{i=1}^{r} \hat{\theta}_i = \theta$.

وفي غير ذلك يقال ان $\hat{\theta}_n$ هو تقدير متحيز biased estimator . ويقال ان $\hat{\theta}_n$ هو تقدير ذو تحيز موجب اذا كان $E\hat{\theta}_n > \theta$ وانه ذو تحيز سالب اذا كان $E\hat{\theta}_n < \theta$. ان مقدار التحيز في التقدير يمكن قياسه وفق الصيغة التالية :

$$\text{bias}(\theta) = E\hat{\theta}_n - \theta \begin{cases} > 0 & \text{تحيز موجب} \\ = 0 & \text{غير متحيز} \\ < 0 & \text{تحيز سالب} \end{cases}$$

مثال (٤) : افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$. بين ان كل من $\bar{x}_n$, $g = \frac{1}{2}(x_1 + x_2)$ هو تقدير غير متحيز الى μ .

الحل :

$$E\bar{x}_n = \frac{1}{n} E \sum_{i=1}^{n} x_i = \frac{1}{n} \sum_{i=1}^{n} Ex_i$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mu = \frac{1}{n} . n\mu = \mu$$

عليه فان $\bar{x}_n$ تقدير غير متحيز الى μ .
كذلك فان

$$Eg = \frac{1}{2} E(x_1 + x_2) = \frac{1}{2}(\mu + \mu) = \mu$$

وهذا يعني ان g تقدير غير متحيز الى μ
نستخلص من المثال (٤) وبشكل عام ان التقدير الغير متحيز لا يتصف بصفة الوحدانية بسبب وجود عدد غير منته من التقديرات الغير متحيزة للمعلمة θ .

مثال (٥) : افرض ان $x_1, x_2, ..., x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$ وافرض ان S_2^2 , S_1^2 يمثلان تقديرين الى σ^2 بحيث ان $S_1^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$ وان $S_2^2 = \frac{1}{n} \Sigma (x_i - \bar{x})^2$ بين ان S_1^2 هو تقدير غير متحيز الى σ^2 في حين ان S_2^2 هو تقدير متحيز الى σ^2 . جد مقدار تحيز هذا التقدير .

الحل :

لاحظنا لدى دراستنا لتوزيع مربع كاي ان

$$Z = \frac{(n-1) S^2}{\sigma^2} \sim \chi^2_{(n-1)} \quad , \quad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

وان

$$EZ = n - 1$$

عليه فان

$$Z_1 = \frac{(n-1) S_1^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$\therefore EZ_1 = \frac{n-1}{\sigma^2} ES_1^2 = n - 1 \quad \therefore ES_1^2 = \sigma^2$$

وهذا يعني ان S_1^2 تقدير غير متحيز الى σ^2 .

وان

$$Z_2 = \frac{nS_2^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$\therefore EZ_2 = \frac{n}{\sigma^2} ES_2^2 = n - 1 \quad \therefore ES_2^2 = \frac{n-1}{n} \sigma^2$$

وهذا يعني ان S_2^2 تقدير متحيز الى σ^2 وان مقدار التحيز هو :

$$bias(\sigma^2) = ES_2^2 - \sigma^2$$

$$= \frac{n-1}{n} \sigma^2 - \sigma^2 = - \frac{\sigma^2}{n}$$

١١ ـ ١ ـ ٣ : الكفاءة Efficiency

لاحظنا من خلال الفقرتين السابقتين ان هنالك اكثر من تقدير واحد يمكن الحصول عليه من العينة يتمتع في ذات الوقت بصفتي الاتساق وعدم التحيز الامر الذي يؤدي الى صعوبة المفاضلة بين هذه التقديرات لاختيار الافضل من بينها. عليه لا بد من ايجاد معيار آخر للمفاضلة. هذا المعيار هو الكفاءة الذي يعتمد بالاساس المفاضلة بين تباين تقديرات من هذا النوع. وتعرف صفة الكفاءة على النحو التالي: بفرض ان θ_2 , θ_1 تقديران متسقان وان $E\hat{\theta}_1 = E\hat{\theta}_2 = \theta$ فـاذا كـان تباين $\hat{\theta}_1$ اقل من تباين $\hat{\theta}_2$ عندئذٍ يقال ان $\hat{\theta}_1$ هو تقدير اكثر كفاءة من $\hat{\theta}_2$ لجميع قيم n. وعلى هذا الاساس يمكن حساب مايسمى بمعامل الكفاءة الذي يمثل النسبة ما بين تباين التقدير الاكثر كفاءة وتباين اي تقدير آخر مثل $\hat{\theta}_2$. فاذا رمزنا لمعامل الكفاءة بالرمز e فان

$$e = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)}$$

ويتضح ان قيمة e هي اقل من الواحد بسبب ان بسط هذه النسبة اقل من مقامها. كذلك فان كفاءة $\hat{\theta}_1$ تزداد بانخفاض قيمة e.

مثال (٦) : اذا علمت ان $\bar{x}_n$ يمثل الوسط الحسابي لقياسات عينة عشوائية ذات حجم n مسحوبة من $N(\mu, \sigma^2)$ وان M_n يمثل الوسيط لهذه العينة وبفرض ان كل من $M_n , \bar{x}_n$ تقدير متسق وغير متحيز الى μ وان

$$V(\bar{x}_n) = \frac{\sigma^2}{n} , V(M_n) = \frac{\pi\sigma^2}{2n}$$

بين ان $\bar{x}_n$ هو تقدير اكثر كفاءة من M_n ثم جد معامل الكفاءة.

الحل :

$$V(M_n) = \frac{\pi\sigma^2}{2n} = 1{\cdot}57V(\bar{x}_n) > V(\bar{x}_n)$$

وهذا يعني ان $\bar{x}_n$ اكثر كفاءة من M_n . وان :

$$e = \frac{V(\bar{x}_n)}{V(M_n)} = 0{\cdot}637$$

مثال (٧) : افرض ان x_1, x_2, x_3, x_4 تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$ وافرض ان كل من $\hat{\theta}_2, \hat{\theta}_1$ تقدير الى μ بحيث ان $\hat{\theta}_1 = \frac{1}{4}\sum_{i=1}^{4} x_i$, $\hat{\theta}_2 = \frac{3x_1 + x_2}{2} - x_3$. بين ان $\hat{\theta}_2, \hat{\theta}_1$ هو تقدير غير متحيز الى μ ثم حدد التقدير الاكثر كفاءة .

الحل :

$$E\hat{\theta}_1 = \frac{1}{4}\sum_{i=1}^{4} Ex_i = \frac{1}{4} . 4\mu = \mu$$

$$E\hat{\theta}_2 = \frac{1}{2}E(3x_1 + x_2) - Ex_3 = 2\mu - \mu = \mu$$

وهذا يعني ان كل من $\hat{\theta}_2, \hat{\theta}_1$ هو تقدير غير متحيز الى μ .

$$V(\hat{\theta}_1) = \frac{1}{16} V\left(\sum_{i=1}^{4} x_i\right) = \frac{1}{16} . 4\sigma^2 = \frac{1}{4}\sigma^2$$

$$V(\hat{\theta}_2) = \frac{1}{4}(9\sigma^2 + \sigma^2) + \sigma^2 = \frac{14}{4}\sigma^2$$

واضح ان $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ وهذا يعني ان $\hat{\theta}_1$ هو التقدير الاكثر كفاءة من $\hat{\theta}_2$ وان :

$$e = \frac{\sigma^2 / 4}{14\sigma^2 / 4} = 0{\cdot}071$$

١١ ـ ١ ـ ٤ : التقدير غير المتحيز ذو أقل تباين . Minimum variance unbiased estimator (M. V. U. E)

يقال إن التقدير $\hat{\theta}_n$ هو تقدير غير متحيز ذو اقل تباين اذا تحقق مايلي :
أ ـ ان $\hat{\theta}_n$ تقدير غير متحيز الى θ . أي ان $E\hat{\theta}_n = \theta$.
ب ـ ان $\hat{\theta}_n$ يمتلك اقل تباين من بين تباينات جملة تقديرات اخرى غير متحيزة الى θ .

مبرهنة : ان التقدير غير المتحيز ذو اقل تباين هو تقدير وحيد .

البرهان :
ليكن t_2, t_1 تقديران غير متحيزان وبأقل تباين الى θ . ان المطلوب برهنته هنا هو ان $t_1 = t_2$

واضح ان $Et_1 = Et_2 = \theta$ وان $V(t_1) = V(t_2)$.

ليكن $t = \frac{1}{2}(t_1 + t_2)$ تقدير آخر الى θ . واضح ان t تقدير غير متحيز الى θ طالما ان :

$$Et = \frac{1}{2}(Et_1 + Et_2) = \theta$$

كذلك فان :

$$V(t) = \frac{1}{4}[V(t_1) + V(t_2) + 2Cov(t_1, t_2)]$$

وحيث ان

$$Cov(t_1, t_2) = \rho\sqrt{V(t_1) . V(t_2)}$$

حيث ان ρ تعني معامل الارتباط البسيط بين t_2, t_1 . لذا فان :

$$V(t) = \frac{1}{4}\left[V(t_1) + V(t_2) + 2\rho\sqrt{V(t_1).V(t_2)}\right]$$

وبجعل $V(t_1) = V(t_2)$ نحصل على :

$$V(t) = \frac{1}{4}\left[2V(t_1) + 2\rho.V(t_1)\right] = \frac{V(t_1)}{2}(1+\rho)$$

الان وحيث ان t_1 يمثل تقدير غير متحيز ذو اقل تباين الى θ فذلك يعني ان :

$$V(t_1) \leq V(t) \rightarrow V(t_1) \leq \frac{V(t_1)}{2}(1+\rho)$$

$$\therefore 1 \leq \frac{1+\rho}{2} \rightarrow \rho \geq 1$$

وكما هو معلوم فان $|\rho| \leq 1$ وهذا يعني ان $\rho = 1$ أي ان t_2, t_1 مرتبطان بعلاقة خطية تامة من الشكل $t_1 = a + bt_2$ حيث ان b, a ثوابت حقيقية لا تعتمد على قياسات العينة وانما قد تعتمد على قيمة المعلمة θ . الان :

$$Et_1 = a + bEt_2 \rightarrow \theta = a + b\theta \dots (*)$$
$$V(t_1) = V(a + bt_2) = b^2V(t_2)$$

لكن $V(t_1) = V(t_2)$. عليه فان :

$$V(t_1) = b^2v(t_1) \rightarrow 1 = b^2 \rightarrow b = \pm 1$$

لكن طالما ان $\rho = 1$ فذلك يعني ان t_2, t_1 مرتبطين بعلاقة خطية تامة موجبة أي ان b يجب ان تكون مساوية للواحد .

وبالتعويض عن $b = 1$ في $(*)$ نحصل على $\theta = a + \theta$ وهذا يعني ان $a = 0$ ، وبالتعويض عن $a = 0$ ، $b = 1$ في العلاقة $t_1 = a + bt_2$ نحصل على $t_1 = t_2$ وهو المطلوب اثباته .

١١ ـ ١ ـ ٥ : الكفاية Sufficiency

يقال ان المؤشر الاحصائي t_n هو تقدير كاف للمعلمة θ اذا كان هذا المؤشر يحوي كافة المعلومات المستقاة من العينة فيما يخص المعلمة θ .
ورياضياً تعرف صفة الكفاية على النحو التالي : بفرض ان المؤشر t_n يمثل دالة بدلالة قياسات عينة عشوائية ذات حجم n مسحوبة من $f(x, \theta)$. فاذا كان التوزيع الشرطي الى $x_1, x_2, \ldots, x_n$ علماً ان t_n معطاة مستقل عن θ عندئذ يقال ان t_n هو مؤشر كاف للمعلمة θ .

مثال (٨) : افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع برنولي بالمعلمة p ، وافرض ان $T_n = \Sigma x_i$ بين ان T_n مؤشر كافٍ للمعلمة P .

الحل :

ان $P(x, p) = P^x (1 - p)^{1-x}, x = 0,1$ وان التوزيع الاحتمالي الى T_n وباستخدام اسلوب الدالة المولدة للعزوم هو الاتي : بفرض $M_{T_n}(t)$ موجودة .
فاذن

$$M_{T_n}(t) = Ee^{t \Sigma x_i}$$

$$= \prod_{i=1}^{n} M_{x_i}(t) = \prod_{i=1}^{n} (q + pe^t) = (q + pe^t)^n$$

وهذا يعني ان $T_n \sim b(n, p)$ فاذن

الان $$P(T_n = k) = C_k^n P^k (1 - P)^{n-k}$$

$$P(x_1, x_2, \ldots, x_n | T_n = k) = \frac{P(x_1, x_2, \ldots, x_n)}{P(T_n = k)}$$

$$= \frac{\prod_{i=1}^{n} P^{x_i} (1 - p)^{1 - x_i}}{C_k^n P^k (1 - p)^{n-k}}$$

$$= \frac{P^{\Sigma x_i} . (1-p)^{n-\Sigma x_i}}{C_k^n P^k (1-p)^{n-k}}$$

$$= \frac{P^k (1-p)^{n-k}}{C_k^n P^k (1-p)^{n-k}} = \frac{1}{C_k^n}$$

وحيث ان التوزيع الشرْطي لايعتمد على p فذلك يعني ان $T_n = \Sigma x_i$ هو مؤشر كاف الى p .

مثال (٩) : افرض ان $x_1, x_2, \dots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع اسي بالمعلمة θ . بين ان المؤشر $T_n = \Sigma x_i$ كاف للمعلمة θ .

الحل : ان

$$f(x, \theta) = \theta e^{-\theta x}$$

ان التوزيع الاحتمالي الى T_n وباستخدام اسلوب الدالة المولدة للعزوم هو الاتي :

بفرض ان $M_{T_n}(t)$ موجودة ، وذلك يعني ان :

$$M_{T_n}(t) = Ee^{t \Sigma x_i} = \prod_{i=1}^{n} M_{x_i}(t)$$

$$= \prod_{i=1}^{n} \left(\frac{\theta}{\theta - t} \right) = \left(1 - \frac{1}{\theta} t \right)^{-n}$$

وهذا يعني ان $T_n \sim G\left(n, \frac{1}{\theta}\right)$ فاذن :

$$f(T_n) = \frac{1}{\Gamma(n) . \theta^{-n}} (T_n)^{n-1} . e^{-\theta T_n}$$

الان

$$f(x_1, x_2, \ldots, x_n \mid T_n = \Sigma x_i) = \frac{f(x_1, x_2, \ldots, x_n)}{f(T_n)}$$

$$= \frac{\prod_{i=1}^{n} \theta e^{-\theta x_i}}{f(T_n)} = \frac{\theta^n . e^{-\theta T_n}}{\frac{\theta^n}{\Gamma(n)} (T_n)^{n-1} . e^{-\theta T_n}}$$

$$= \frac{\Gamma(n)}{(T_n)^{n-1}}$$

وحيث ان التوزيع الشرطي لا يعتمد على θ فذلك يعني ان $T_n = \Sigma x_i$ هو مؤشر كاف للمعلمة θ .

مبرهنة التحليل Factorization theorem

ان هذه المبرهنة مهمة جداً في الحصول على مؤشرات كافية . وتنص هذه المبرهنة بما يلي :

افرض ان المؤشر الاحصائي T_n دالة بدلالة قياسات العينة . أي ان $T_n = T(x_1, x_2, \ldots, x_n)$ عندئذٍ يقال ان T_n مؤشر كاف للمعلمة θ اذا وفقط اذا امكن تحليل الدالة المشتركة لقياسات العينة الى حاصل ضرب مركبتين الاولى تمثل دالة بدلالة T_n . θ والثانية دالة لاتعتمد على θ . وبالرموز وبفرض ان L تمثل الدالة المشتركة لقياسات العينة وان $g(T_n, \theta)$ تمثل دالة بدلالة المؤشر T_n والمعلمة θ وان $h(x)$ تمثل داله لا تعتمد على θ عندئذٍ يقال ان T_n مؤشر كاف الى θ اذا وفقط اذا كان :

$$L = g(T_n, \theta) . h(x)$$

مثال (١٠) : افرض ان $x_1, x_2, \ldots, x_n$. تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع پواسون بالمعلمة λ . بين ان Σx_i هو مؤشر كاف للمعلمة λ .

الحل

$$L = P(x_1, x_2, \ldots, x_n, \lambda) = \prod_{i=1}^{n} P(x_i, \lambda)$$

$$= \prod_{i=1}^{n} \frac{\lambda^{x_i} . e^{-\lambda}}{x_i !}$$

$$= \frac{\lambda^{\Sigma x_i} . e^{-n\lambda}}{\prod_{i=1}^{n} x_i !} = (\lambda^{\Sigma x_i} . e^{-n\lambda}) . \frac{1}{\prod_{i=1}^{n} x_i !}$$

يتضح ان

$$h(x) = \frac{1}{\prod_{i=1} x_i !}, g(\Sigma x_\theta, \lambda) = \lambda^{\Sigma x_i} . e^{-n\lambda}$$

وهذا يعني انه حسب مبرهنة التحليل ان الموشر الاحصائي Σx_i كاف للمعلمة λ .

مثال (١١) : لتكن $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع بيتا بالمعلمتين $\alpha, 2 = \beta$ بين ان $\prod_{i=1}^{n} x_i$ هو مؤشر كاف للمعلمة α

الحل : ان

$$f(x, \alpha, \beta = 2) = \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)\Gamma(2)} x^{\alpha-1}(1 - x) ; 0 < x < 1$$

$$= \alpha(\alpha + 1) x^{\alpha-1} . (1 - x)$$

$$\therefore L = \prod_{i=1}^{n} f(x_i, \alpha, \beta = 2) = \alpha^n (\alpha + 1)^n . \left(\prod_{i=1}^{n} x_i \right)^{\alpha - 1} \prod_{i=1}^{n} (1 - x_i)$$

واضح هنا ان

$$g(T_n, \theta) = g\left(\prod_{i=1}^{n} x_i , \alpha \right)$$

$$= \alpha^n (\alpha + 1)^n . \left(\prod_{i=1}^{n} x_i \right)^{\alpha - 1}$$

وان

$$h(x) = \prod_{i=1}^{n} (1 - x_i)$$

وهذا يعني ان $\prod_{i=1}^{n} x_i$ مؤشر كاف للمعلمة α

١١ ـ ٢ : طرق التقدير Methods of estimation

تطرقنا في محتويات الفقرة (١١ ـ ١) الى استعراض لاهم الصفات التي يجب ان يتمتع بها التقدير كي يسمح لنا ذلك تسمية ذلك التقدير « تقدير جيد » : في هذه الفقرة سوف نتطرق وبشكل موجز الى اهم الطرق التي تقودنا للحصول على تقديرات جيدة لمعلمة او معالم مجتمع احصائـي معين . هذه الطرق هي :

١ ـ طريقة الامكان الاعظم .
٢ ـ طريقة التباين الاقل .
٣ ـ طريقة المربعات الصغرى .
٤ ـ طريقة العزوم .
٥ ـ طريقة اقل مربع كاي ممكن .
٦ ـ طريقة الاحتمال المعكوس .

وسوف نستعرض وبشكل موجز الطريقتين (١) ، (٢) فقط كي لا نخرج عن الهدف المتوخى من هذا الكتاب .

١١ ـ ٢ ـ ١ : طريقة الامكان الاعظم Method of maximum likelihood

تعد طريقة الامكان الاعظم احدى اهم طرق التقدير . وقبيل الحديث عن هذه الطريقة لابد من اعطاء تعريف لما هو مقصود بـ « دالة الامكان likelihood function». لتكن $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع بدالة كتلة احتمالية $p(x, \theta)$ او دالة كثافة احتمالية $f(x, \theta)$. عندئذ تعرف دالة الامكان لقياسات العينة بانها التوزيع المشترك لتلك القياسات . فاذا رمزنا لدالة الامكان بالرمز L فان هذه الدالة في حالة المتغيرات المتقطعة هي :

$$L = p(x_1, x_2, \ldots, x_n, \theta) = \prod_{i=1}^{n} P(x_i, \theta)$$

وفي حالة المتغيرات المستمرة فان :

$$L = f(x_1, x_2, \ldots, x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

ان مبدأ طريقة الامكان الاعظم يكمن في ايجاد تقدير مثل $\hat{\theta}$ للمعلمة θ الذي يجعل دالة الامكان L في نهايتها العظمى . فاذا كان $\hat{\theta}$ دالة بدلالة قياسات العينة التي تجعل L في نهايتها العظمى عندئذٍ يقال ان $\hat{\theta}$ هو تقدير الامكان الاعظم (M.L.E) الى θ . وهذا يعني ان $\hat{\theta}$ ناتج من حل المعادلة التفاضلية :

$$\frac{\partial L}{\partial \theta} = 0 \quad \text{بشرط ان} \quad \frac{\partial^2 L}{\partial \theta^2} < 0$$

وبهدف السهولة في اجراء عمليات التفاضل السابقة فانه غالباً ما يتم التعامل مع Log L . حيث ان $L > 0$ فان :

$\frac{\partial L}{\partial \theta} = 0$ مكافئة الى $\frac{\partial \operatorname{Log} L}{\partial \theta} = 0$ وهذا يعني ان :

$$\frac{\partial \operatorname{Log} L}{\partial \theta} = \frac{1}{L} \cdot \frac{\partial L}{\partial \theta} = 0$$

علماً ان مبدأ طريقة الامكان الاعظم يمكن تطبيقه في حالة التوزيعات متعددة المتغيرات والتوزيعات ذات معلمتين او اكثر الا اننا سنكتفي بعرض حالة التوزيعات ذات المتغير الواحد والاتي بعض خصائص تقديرات الامكان الاعظم ندرجها ادناه دون اللجوء الى تقديم البراهين اللازمة لها وهي :

١ ـ ان تقديرات الامكان الاعظم هي تقديرات متسقة .
٢ ـ ان التوزيع الاحتمالي لتقدير الامكان الاعظم يؤول الى التوزيع الطبيعي عندما يكون حجم العينة كبير (نظرياً ∞) .
٣ ـ ان تقدير الامكان الاعظم هو تقدير اكثر كفاءة من بين جملة تقديرات اخرى متاحة .
٤ ـ اذا كان T_n موشر كاف للمعلمة θ فان تقدير الامكان الاعظم $\hat{\hat{\theta}}$ يكون دالة بدلالة T_n
٥ ـ ان تقدير الامكان الاعظم $\hat{\theta}$ ليس بالضرورة ان يكون تقدير غير متحيز الى θ .
٦ ـ ان التباين المحاذي Asymptotic variance لتقدير الامكان الاعظم $\hat{\theta}$ معطى بالصيغة التالية :

$$V(\hat{\theta}) = \left[E\left\{ -\left(\frac{\partial^2 \log L}{\partial \theta^2} \right) \right\} \right]^{-1}$$

وان $V(\hat{\theta}) \leq V(\hat{\hat{\theta}})$ حيث ان $\hat{\hat{\theta}}$ يمثل اي تقدير آخر الى θ .

مثال (١٢) : لتكن $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع اسي بالمعلمة θ . جد تقدير الامكان الاعظم الى θ . ثم جد تباين هذا التقدير .

الحل :

$$L = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n . e^{-\theta \Sigma x_i}$$

فاذن

$$\text{Log } L = n \log \theta - \theta \Sigma x_i$$

عليه فان

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} - \Sigma x_i , \quad \frac{\partial^2 \log L}{\partial \theta^2} = - \frac{n}{\theta^2} < 0$$

وبجعل $\frac{\partial \text{Log } L}{\partial \theta} = 0$ فان :

$$\frac{n}{\theta} - \Sigma x_i = 0 \rightarrow \frac{n}{\theta} = \Sigma x_i$$

وهذا يعني ان تقدير الامكان الاعظم الى θ هو

$$\hat{\theta} = \frac{n}{\Sigma x_i} = \frac{1}{\bar{x}}$$

وان

$$V(\hat{\theta}) = \left[E \left\{ - \left(- \frac{n}{\theta^2} \right) \right\} \right]^{-1} = \left[\frac{n}{\theta^2} \right]^{-1} = \frac{\theta^2}{n}$$

مثال (١٣) : افــرض ان $x_1 , x_2 , \dots , x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع پواسون بالمعلمة λ . جد تقدير الامكان الاعظم الى λ ثم جد تباين هذا التقدير .

الحل :

$$L = \prod_{i=1}^{n} P(x_i , \lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} . e^{-\lambda}}{x_i !} = \frac{\lambda^{\Sigma x_i} \ e^{-n\lambda}}{\prod_{i=1}^{n} x_i !}$$

$$\therefore \text{Log } L = \Sigma x_i \log \lambda - n\lambda - \Sigma \log x_i !$$

$$\therefore \frac{\partial \log L}{\partial \lambda} = \frac{\Sigma x_i}{\lambda} - n , \quad \frac{\partial^2 \log L}{\partial \lambda^2} = - \frac{\Sigma x_i}{\lambda^2} < 0$$

وبجعل $\frac{\partial \log L}{\partial \lambda} = 0$ فان :

$$\frac{\Sigma x_i}{\lambda} - n = 0 \rightarrow \hat{\lambda} = \frac{\Sigma x_i}{n} = \bar{x}$$

وهذا يعني ان تقدير الامكان الاعظم الى λ هو متوسط قياسات العينة وان :

$$V(\lambda) = \left[E\left\{ -\left(-\frac{n\bar{x}}{\lambda^2} \right) \right\} \right]^{-1}$$

$$= \left[\frac{n}{\lambda^2} E\bar{x} \right]^{-1} = \left[\frac{n}{\lambda} \right]^{-1} = \frac{\lambda}{n}$$

مثال (١٤) : لتكن $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$. جد تقدير الامكان الاعظم الى :

أ ـ μ بفرض ان σ^2 معلومة . ثم جد تباين هذا التقدير .

ب ـ σ^2 بفرض ان μ معلومة . ثم جد تباين هذا التقدير .

جـ ـ μ, σ^2 معاً بفرض ان كليهما مجهولين .

الحل :

$$L = \prod_{i=1}^{n} . e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2}$$

$$= \left(\frac{1}{\sqrt{2\pi}\ \sigma} \right)^n . e^{-\frac{1}{2\sigma^2}\Sigma(x_i - \mu)^2}$$

$$\therefore \log L = n \log \left(\frac{1}{\sqrt{2\pi}} \right) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \Sigma(x_i - \mu)^2$$

أ ـ اذا كانت σ^2 معلومة فان :

$$\frac{\partial \log L}{\partial \mu} = \frac{1}{\sigma^2} \Sigma (x_i - \mu) , \quad \frac{\partial^2 \log L}{\partial \mu^2} = - \frac{n}{\sigma^2} < 0$$

وبجعل $\frac{\partial \log L}{\partial \mu} = 0$ فان :

$$\frac{1}{\sigma^2} \Sigma (x_i - \mu) = 0 \rightarrow \hat{\mu} = \bar{x}$$

أي ان تقدير الامكان الاعظم الى μ بفرض ان σ^2 معلومة هو متوسط قياسات العينة . وان :

$$V (\mu) = \left[E \left\{ - \left(- \frac{n}{\sigma^2} \right) \right\} \right]^{-1} = \left[\frac{n}{\sigma^2} \right]^{-1} = \frac{\sigma^2}{n}$$

ب ـ اذا كانت μ معلومة فان :

$$\frac{\partial \log L}{\partial \sigma^2} = - \frac{n}{2 \sigma^2} + \frac{1}{2 \sigma^4} \Sigma (x_i - \mu)^2$$

وان

$$\frac{\partial^2 \log L}{\partial (\sigma^2)^2} = \frac{n}{2 \sigma^4} - \frac{1}{\sigma^6} \Sigma (x_i - \mu)^2$$

$$= \frac{n}{2\sigma^4} - \frac{1}{\sigma^4} \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^2$$

لكن

$$\sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^2 \sim \chi^2_{(n)}$$

فاذن

$$\frac{\partial^2 \log L}{\partial(\sigma^2)^2} = \frac{1}{\sigma^4}\left[\frac{n}{2} - \chi^2_{(n)}\right] < 0 \,,\; E\,\chi^2_{(n)} = n$$

عليه وبجعل $0 = \frac{\partial \log L}{\partial \sigma^2}$ فأن

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\Sigma(x_i - \mu)^2 = 0$$

أو ان

$$\frac{1}{\sigma^2}\Sigma(x_i - \mu)^2 = n$$

عليه فان تقدير الامكان الاعظم هو

$$\hat{\sigma}^2 = \frac{1}{n}\Sigma(x_i - \mu)^2$$

كذلك يلاحظ هنا ان

$$\frac{n\hat{\sigma}^2}{\sigma^2} = \Sigma\left(\frac{x_i - \mu}{\sigma}\right)^2 \sim \chi^2_{(n)}$$

$$\therefore\ \frac{n}{\sigma^2}E\hat{\sigma}^2 = n \rightarrow E\hat{\sigma}^2 = \sigma^2$$

أي ان تقدير الامكان الاعظم عندما μ معلومة هو تقدير غير متحيز الى σ^2
وان تباين هذا التقدير هو :

$$V(\hat{\sigma}^2) = \left[E\left\{\frac{1}{\sigma^4}\left(\chi^2_{(n)} - \frac{n}{2}\right)\right\}\right]^{-1}$$

$$= \left[\frac{1}{\sigma^4}\left(n - \frac{n}{2}\right)\right]^{-1} = \left[\frac{n}{2\sigma^4}\right]^{-1} = \frac{2\sigma^4}{n}$$

جـ ـ لاحظنا ان حل المعادلة التفاضلية $\frac{\partial \log L}{\partial \mu} = 0$ كان $\hat{\mu} = \bar{x}$. عليه فان حل المعادلة التفاضلية $\frac{\partial \log L}{\partial \sigma^2} = 0$ عندما $\hat{\mu} = \bar{x}$ هو

$$\hat{\sigma}^2 = \frac{1}{n} \Sigma (x_i - \bar{x})^2 = S^2$$

أي ان تقدير الامكان الاعظم الى σ^2 عندما μ مجهولة هو تباين العينة S^2 . ويلاحظ هنا ان S^2 هو تقدير متحيز الى σ^2 كما سبق وان اوضحناه في المثال (٥) من هذا الفصل لدى دراستنا لخاصية عدم التحيز .

١١ ـ ٢ ـ ٢ : طريقة التباين الاقل Method of minimum variance

يتم وفق هذه الطريقة ايجاد تقديرات غير متحيزة ذات اقل تباين minimum variance unbiased estimators (M. V. U. E) ويستند مبدأ هذه الطريقة على ما يلي :

افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع بدالة كتلة احتمالية $p(x, \theta)$ او دالة كثافة احتمالية $f(x,\theta)$ معلمة مجهولة ، وافرض ان L تمثل دالة الامكان وان $t = t(x_1, x_2, \ldots, x_n)$ تمثل دالة بدلالة قياسات العينة خالية من اي مجهول وان λ تمثل دالة بدلالة θ مستقلة عن قياسات العينة . عندئذٍ فان الشرط الكافي والضروري لوجود تقدير غير متحيز ذو اقل تباين مثل t هو امكانية صياغة المشتقة الجزئية الاولى لدالة الامكان بالشكل التالي .

$$\frac{\partial \log L}{\partial \theta} = \frac{t - \theta}{\lambda}$$

وعندئذٍ يقال ان t هو تقدير غير متحيز ذو اقل تباين للمعلمة θ وان λ تمثل تباين هذا التقدير .

مثال (١٥) افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$ ، σ^2 معلومة . جد التقدير الغير متحيز ذو اقل تباين الى μ .

الحل :

$$L = \prod_{i=1}^{n} f(x_i, \mu, \sigma^2) = (\sqrt{2\pi}\ \sigma)^{-n} . e^{-\frac{1}{2\sigma^2}\Sigma(x_i - \mu)^2}$$

$$\therefore \log L = \log(\sqrt{2\pi}\ \sigma)^{-n} - \frac{1}{2\sigma^2}\Sigma(x_i - \mu)^2$$

$$\therefore \frac{\partial \log L}{\partial \mu} = \frac{1}{\sigma^2}\Sigma(x_i - \mu) = \frac{n\bar{x} - n\mu}{\sigma^2} = \frac{\bar{x} - \mu}{\sigma^2 / n}$$

لاحظ هنا ان $\lambda = V(\bar{x}) = \frac{\sigma^2}{n}, \theta = \mu, t = \bar{x}$. فاذن $\bar{x}$ هو التقدير غير المتحيز ذو اقل تباين للمعلمة μ .

مثال (١٦) : لتكن $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $b(n,p)$. جد التقدير الغير متحيز ذو أقل تباين للمعلمة p .

الحل :

$$L = \prod_{i=1}^{n} C^n_{x_i}\ p^{x_i} . (1 - p)^{n - x_i}$$

$$= \left(\prod_{i=1}^{n} C^n_{x_i}\right) . p^{\Sigma x_i} . (1 - p)^{n^2 - \Sigma x_i}$$

$$\therefore \log L = \log\left(\prod_{i=1}^{n} C^n_{x_i}\right) + \Sigma x_i \log p + (n^2 - \Sigma x_i) \log(1 - p)$$

$$\therefore \frac{\partial \log L}{\partial p} = \frac{\Sigma x_i}{p} - \frac{n^2 - \Sigma x_i}{1 - p}$$

$$= \frac{n\bar{x} - n^2 p}{p(1-p)}, \quad n\bar{x} = \Sigma x_i$$

$$= \frac{\frac{\bar{x}}{n} - p}{\frac{p(1-p)}{n^2}}$$

لاحظ هنا ان $t = \frac{\bar{x}}{n}$, $\theta = p$, $\lambda = V\left(\frac{\bar{x}}{n}\right) = \frac{pq}{n^2}$. فاذن $\frac{\bar{x}}{n}$ هو التقدير غير المتحيز ذات اقل تباين الى p .

١١ ــ ٣ : التقدير بفترة Interval estimation

استعرضنا في الفقرتين السابقتين اهم الصفات التي يجب ان يتمتع بها التقدير لمعلمة معينة وكذلك استعرضنا اهم طريقتين يمكن من خلالهما التوصل الى التقدير الجيد .

في بعض الاحيان قد لا نرغب في ايجاد تقدير (قيمة واحدة فقط) لمعلمة مجهولة مثل θ وانما نرغب في ايجاد فترة يتوقع ان نلاحظ تلك المعلمة خلالها بدرجة ثقة معينة . وهذا ما نسميه التقدير بفترة وان تلك الفترة التي نحصل عليها تسمى فترة ثقة confidence interval .

الان وبفرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع بدالة كتلة احتمالية $P(x, \theta)$ أو دالة كثافة احتمالية $f(x, \theta)$ وان $t = t(x_1, x_2, \ldots, x_n)$ تمثل دالة بدلالة قياسات العينة خالية من اي مجهول وان t تمثل تقدير للمعلمة θ . ليكن $g(t \mid \theta)$ يمثل

توزيع المعاينة للمؤشر الاحصائي t وان C_2 , C_1 ثابتان حقيقيان يتحددان من خلال موشرات العينة (كالوسط والتباين مثلاً) وتوزيع المعاينة للمؤشر t . عندئذٍ فان مهمة التقدير بفترة هي حساب C_2 , C_1 التي تقودنا الى صياغة الجملة الاحتمالية التالية :

$$P_r (C_1 \leq \theta \leq C_2 \mid t) = 1 - \alpha , 0 < \alpha < 1$$

حيث ان α عدد صغير (عمليا تختار α لان تكون % 5 أو % 1 ويعتمد هذا الاختيار بطبيعة الحال على درجة الثقة المطلوبة للفترة) وان $(1 - \alpha)$ يسمى معامل الثقة . ان هذه الجملة الاحتمالية مفادها الاتي : ان احتمال ملاحظة θ في الفترة (C_1 , C_2) بالاستناد لمعطيات العينة وتوزيع المعاينة للمؤشر t هو $(1 - \alpha)$ حيث ان C_2 , C_1 يسميان حدي الثقة confidence limits . وسوف نركز اهتمامنا هنا على حساب فترات الثقة لمعالم مجتمع طبيعي وخصوصاً ما يتعلق الامر بالوسط والتباين .

١١ ــ ٣ ــ ١ : فترة ثقة لمتوسط مجتمع طبيعي

افرض ان $x_1 , x_2 , \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N (\mu , \sigma^2)$ وان $\bar{x}$ يمثل الوسط الحسابي لقياسات هذه العينة . وافرض اننا نرغب في ايجاد فترة ثقة للمعلمة μ .

١ ــ اذا كانت σ^2 معلومة :

كما هو معلوم فانه اذا كان $X \sim N (\mu , \sigma^2)$ فان توزيع $\bar{x}$ سيكون $N \left(\mu, \frac{\sigma^2}{n} \right)$. وعندئذٍ فان :

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N (0.1)$$

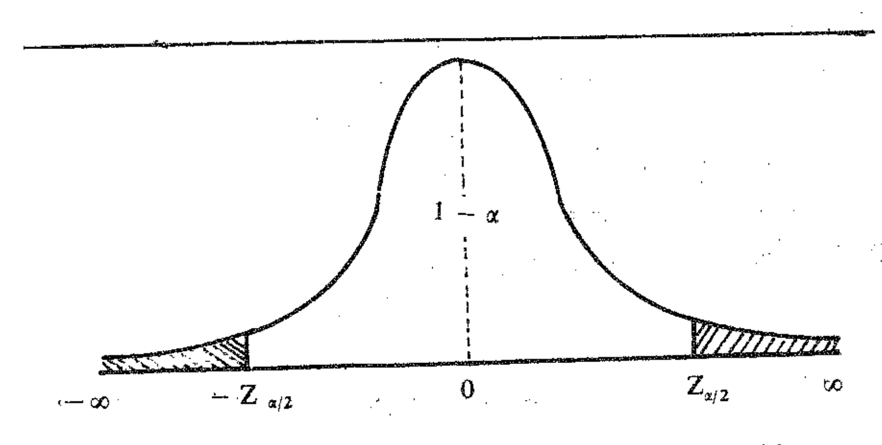
وبذلك فان

$$P_r (| Z | < Z_{\frac{\alpha}{2}}) = P_r (- Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1 - \alpha$$

حيث ان

$$\int_{-\infty}^{-Z_{\frac{\alpha}{2}}} f(z)\,dz = \int_{Z_{\frac{\alpha}{2}}}^{\infty} f(z)\,dz = \frac{\alpha}{2}$$

وكما هو موضح في الشكل (١١ ــ ١) .



شكل (١١ ــ ١) ، توضح لقيمتي $Z\alpha/_2$ ، $-Z\alpha/_2$

وهذا يعني ان $Z_{\frac{\alpha}{2}}$ تتحدد من جداول التوزيع الطبيعي وتعني قيمة Z النظرية التي تعطي احتمالاً قدره $1 - \frac{\alpha}{2}$ للفترة $(-\infty , Z_{\frac{\alpha}{2}}]$ او احتمالاً قدره $\frac{\alpha}{2}$ للفترة $[Z_{\frac{\alpha}{2}} , \infty)$. فمثلاً اذا كانت $\alpha = 0{\cdot}05$ فان

$Z_{0{\cdot}025} = 1{\cdot}96$

الان

$$P_r (-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = P_r \left(-Z_{\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < Z_{\frac{\alpha}{2}} \right)$$

$$= P_r \left(-Z_{\frac{\alpha}{2}} \cdot \frac{\alpha}{\sqrt{n}} < \bar{X} - \mu < Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

$$= P_r \left(-\bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < -\mu < -\bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

$$= P_r\left(\bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} > \mu > \bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

$$= P_r\left(\bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

عليه فان حدي الثقة بمعامل ثقة $(1 - \alpha)$ لمتوسط مجتمع طبيعي تباينه معلوم هما :

$$C_1 = \bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \quad , \quad C_2 = \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

مثال (١٧) : جد حدي الثقة بمعامل ثقة 0.95 لمتوسط مجتمع توزيعه $N(\mu,25)$ علماً ان $n = 100 , \bar{X} = 10$.

الحل :

واضح ان $1 - \alpha = 0.95$. فاذن $\alpha = 0.05$ وان $\frac{\alpha}{2} = 0.025$. من جداول التوزيع الطبيعي فان $Z_{0.025} = 1.96$. عليه فان حدي الثقة لمتوسط هذا المجتمع هما :

$$C_1 = \bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 10 - (1.96) \cdot \frac{5}{10} = 9.02$$

$$C_2 = \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 10 + (1.96) \cdot \frac{5}{10} = 10.98$$

وهذا يعني ان $P_r(9.02 < \mu < 10.98) = 0.95$ اي بثقة مقدارها 0.95 ووفق معطيات هذه العينة يمكن ملاحظة ان μ تقع في الفترة (9.02 , 10.98) .

٢ ـ اذا كانت σ^2 غير معلومة .

اذا كان تباين المجتمع غير معلوم فذلك يعني انه يتوجب تقدير قيمة σ^2 على اساس قياسات تلك العينة . وكما لاحظنا سابقاً فان $S^2 = \frac{1}{n-1}\Sigma(x_i - x)^2$ هو تقدير جيد الى σ^2 . وهنا نقف امام حالتين عند تكوين فترة ثقة الى μ .

أ ـ عندما يكون حجم العينة صغير

في حالة التعامل مع عينات صغيرة (تطبيقياً $n \leq 30$) فان توزيع المتغير العشوائي $t = \frac{\sqrt{n}\,(\bar{X} - \mu)}{S}$ سوف لن يكون $N(0,1)$ وانما سيتوزع كتوزيع t بـ $(n-1)$درجة حرية . أي ان

$$t = \frac{(\bar{X} - \mu)}{S/\sqrt{n}} \sim t_{(n-1)}$$

وعندئذٍ فان

$$P_r(|t| < t_{\frac{\alpha}{2}}) = P_r(-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}}) = 1 - \alpha$$

حيث ان

$$\int_{-\infty}^{-t_{\frac{\alpha}{2}}} g(t)\,dt = \int_{t_{\frac{\alpha}{2}}}^{\infty} g(t)\,dt = \frac{\alpha}{2}$$

وهذا يعني ان $t_{\frac{\alpha}{2}}$ تتحدد من جدول توزيع t وتعني قيمة t النظرية عند درجة حرية $(n-1)$ التي تعطي احتمالاً قدره $\left(1 - \frac{\alpha}{2}\right)$ للفترة $[-\infty, t_{\frac{\alpha}{2}}]$ ، أو احتمالاً قدره $\frac{\alpha}{2}$ للفترة $[t_{\frac{\alpha}{2}}, \infty)$

الان

$$P_r\left(-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}}\right) = P_r\left(-t_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\frac{\alpha}{2}}\right)$$

$$= P_r\left(\bar{X} - t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}\right) = 1-\alpha$$

وبذلك فان حدي الثقة لمتوسط مجتمع طبيعي تباينه مجهول وعند التعامل مع عينات صغيرة $(n \leq 30)$ هما :

$$C_1 = \bar{X} - t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} \ , \ C_2 = \bar{X} + t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$$

مثال (١٨) : افرض ان معطيات عينة عشوائية قوامها 15 مفردة مسحوبة من $N(\mu, \sigma^2)$ كانت $S^2 = 9 , \bar{x} = 20$. جد حدي الثقة بمعامل ثقة 0.95 لمتوسط هذا المجتمع .

الحل :

ان $1 - \alpha = 0.95$ فاذن $\alpha = 0.05$ وان $\frac{\alpha}{2} = 0.025$. من جداول توزيع t عند درجة حرية 14 نلاحظ ان $t_{0.025}(14) = 2.145$ عليه فان

$$C_1 = \bar{X} - t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} = 20 - (2.145) \cdot \frac{3}{\sqrt{15}} = 18.3385$$

$$C_2 = \bar{X} + t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} = 20 + (2.145) \cdot \frac{3}{\sqrt{15}} = 21.6615$$

وهذا يعني ان $P_r(18.3385 < \mu < 21.6615) = 0.95$

ب ـ عندما يكون حجم العينة كبير .

في حالة التعامل مع عينات كبيرة الحجم (تطبيقياً $30 < n$) فان توزيع المتغير العشوائي $t = \frac{\sqrt{n}\,(\bar{X} - \mu)}{S}$ سوف يتقارب من $N(0,1)$. حيث سبق وان برهنا في الفقرة (٩ ـ ٢ ـ ٦) ان :

$$\lim_{n \to \infty} g(t) \to N(0,1)$$

وهذا يعني السماح باستخدام جداول $N(0,1)$ كبديل لجداول توزيع t .. وبذلك فان حدي الثقة لمتوسط المجتمع هما :

$$C_1 = \bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}, \quad C_2 = \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$$

مثال (١٩) : افرض ان معطيات عينة عشوائية ذات حجم 100 مسحوبة من $N(\mu, \sigma^2)$ كانت $x = 20$, $S^2 = 16$. جد حدي الثقة بمعامل ثقة 0.99 لمتوسط هذا المجتمع .

الحل :

واضح ان $1 - \alpha = 0{\cdot}99$ اذن $\alpha = 0{\cdot}01$ وان $\frac{\alpha}{2} = 0{\cdot}005$ من جداول التوزيع الطبيعي نجد ان $Z_{0{\cdot}005} = 2{\cdot}58$. عليه فان :

$$C_1 = \bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} = 20 - (2{\cdot}58) \cdot \frac{4}{10} = 18{\cdot}968$$

$$C_2 = \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} = 20 + (2{\cdot}58) \cdot \frac{4}{10} = 21{\cdot}032$$

وهذا يعني ان

$$P_r(18{\cdot}968 < \mu < 21{\cdot}032) = 0{\cdot}99$$

١١ ـ ٣ ـ ٢ : فترة ثقة لتباين مجتمع طبيعي

افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$. وافرض اننا نرغب في ايجاد فترة ثقة بمعامل ثقة $(1 - \alpha)$ لتباين هذا المجتمع . وهنا نقف امام حالتين :

١ ـ اذا كان متوسط المجتمع μ معلوم .

اذا كانت μ معلومة فان $S^2 = \frac{1}{n} \Sigma (x_i - \mu)^2$ سوف يمثل تباين العينة على اساس متوسط المجتمع μ . وكما هو معلوم فان المتغير العشوائي $\chi^2 = \frac{nS^2}{\sigma^2}$ سوف يتوزع كتوزيع مربع كاي بـ n درجة حرية اي ان :

$$\chi^2 = \frac{nS^2}{\sigma^2} \sim \chi^2_{(n)}$$

وعندئذٍ فان :

$$P_r(\chi^2_{\frac{\alpha}{2}} < \chi^2 < \chi^2_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

حيث ان

$$\int_0^{\chi^2_{\frac{\alpha}{2}}} f(\chi^2)\, d\chi^2 = \frac{\alpha}{2}, \int_0^{\chi^2_{1-\frac{\alpha}{2}}} f(\chi^2)\, d\chi^2 = 1 - \frac{\alpha}{2}$$

وهذا يعني ان $\chi^2_{\frac{\alpha}{2}}$ تمثل قيمة χ^2 النظرية عند درجة حرية n التي تعطي احتمالاً قدره $\frac{\alpha}{2}$ للفترة $(0, \chi^2_{\frac{\alpha}{2}})$ وان $\chi^2_{1-\frac{\alpha}{2}}$ تمثل قيمة χ^2 النظرية عند درجة حرية n التي تعطي احتمالاً قدره $1 - \frac{\alpha}{2}$ للفترة $(0, \chi^2_{1-\frac{\alpha}{2}})$. عليه فان :

$$P_r(\chi^2_{\frac{\alpha}{2}} < \chi^2 < \chi^2_{1-\frac{\alpha}{2}}) = P_r(\chi^2_{\frac{\alpha}{2}} < \frac{nS^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}})$$

$$= P_r\left(\frac{1}{\chi^2_{\frac{\alpha}{2}}} > \frac{\sigma^2}{nS^2} > \frac{1}{\chi^2_{1-\frac{\alpha}{2}}}\right)$$

$$= P_r\left(\frac{nS^2}{\chi^2_{\frac{\alpha}{2}}} > \sigma^2 > \frac{nS^2}{\chi^2_{1-\frac{\alpha}{2}}}\right)$$

$$= P_r\left(\frac{nS^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{nS^2}{\chi^2_{\frac{\alpha}{2}}}\right) = 1 - \alpha$$

عليه فان حدي الثقة هما :

$$C_1 = \frac{nS^2}{\chi^2_{1-\frac{\alpha}{2}}}, \quad C_2 = \frac{nS^2}{\chi^2_{\frac{\alpha}{2}}}$$

مثال (٢٠) : اذا علمت ان تبايـن عينة عشوائية ذات حجم 25 مسحوبة من $N(10, \sigma^2)$ كان $S^2 = \frac{1}{25}\Sigma(x_i - 10)^2 = 9$. جد حدي الثقة بمعامل ثقة 0.95 لتباين هذا المجتمع .

الحل :

ان $1 - \alpha = 0{\cdot}95$. فاذن $\alpha = 0{\cdot}05$ وان $\frac{\alpha}{2} = 0{\cdot}025$. من جداول توزيع مربع كاي بدرجة حرية (25) فان :

$$\chi^2_{0{\cdot}025}(25) = 13{\cdot}1197, \ \chi^2_{0{\cdot}975}(25) = 40{\cdot}6465$$

عليه فان :

$$C_1 = \frac{nS^2}{\chi^2_{1-\frac{\alpha}{2}}} = \frac{(25)(9)}{40{\cdot}6465} = 5{\cdot}5355$$

$$C_2 = \frac{nS^2}{\chi^2_{\frac{\alpha}{2}}} = \frac{(25)(9)}{13{\cdot}1197} = 17{\cdot}1498$$

وهذا يعني ان :

$$P_r(5{\cdot}5355 < \sigma^2 < 17{\cdot}1498) = 0{\cdot}95$$

٢ ـ اذا كان متوسط المجتمع μ مجهول .

في حالة كون متوسط المجتمع مجهول فان افضل تقدير الى تباين هذا المجتمع هو،

$$S^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$$

وفي هذه الحالة فان المتغير العشوائي

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

وعندئذٍ فان حدي الثقة لتباين المجتمع هما ،

$$C_1 = \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} , C_2 = \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}}$$

مثال (٢١) : اجد حدي الثقة لتباين مجتمع طبيعي متوسطه مجهول وفق المعطيات التالية : $1 - \alpha = 0{\cdot}95 , n = 26 , S^2 = 16$.

الحل : ان

$$\chi^2_{0\cdot025}(25) = 13{\cdot}1197 , \chi^2_{0\cdot975}(25) = 40{\cdot}6465$$

عليه فان

$$C_1 = \frac{(25)(16)}{40{\cdot}6465} = 9{\cdot}8409 , C_2 = \frac{(25)(16)}{13{\cdot}1197} = 30{\cdot}4885$$

فاذن

$$P_r(9{\cdot}8409 < \sigma^2 < 30{\cdot}4885) = 0{\cdot}95$$

١١ ــ ١ : افرض ان $x_1, x_2, ..., x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(0, \sigma^2)$. بين ان $\frac{1}{n}\Sigma x_i^2$ هو تقدير غير متحيز الى σ^2 وان تباين هذا التقدير هو $\frac{2\sigma^4}{n}$ ثم بين ان Σx_i^2 هو مؤشر كاف الى σ^2 .

١١ ــ ٢ : لتكن $x_1, x_2, ..., x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع هندسي بالمعلمة P . بين ان Σx_i هو مؤشر كاف الى P . جد تقدير الامكان الاعظم الى P ثم جد تباين هذا التقدير .

١١ ــ ٣ : اذا علمت ان x_1, x_2, x_3 تمثل قياسات عينة عشوائية مسحوبة من $N(\mu, \sigma^2)$ وافرض ان $\hat{\theta}_1$ ، $\hat{\theta}_2$ ، $\hat{\theta}_3$ تمثل ثلاثة تقديرات الى μ معرفة بالشكل التالي :

$$\hat{\theta}_1 = \frac{1}{3}(x_1 + x_2 + x_3), \hat{\theta}_2 = \frac{3x_1 + 5x_2}{2} - 3x_3, \hat{\theta}_3 = 2x_1 - x_2$$

بين ان هذه التقديرات هي تقديرات غير متحيزة الى μ . أي من هذه التقديرات هو الاكفأ .

١١ ــ ٤ : اذا كانت $x_1, x_2, ..., x_n$ تمثل قياسات عينة عشوائية مسحوبة من $b(n, P)$. بين ان Σx_i هو مؤشر كاف للمعلمة P . هل يمكن القول ان $\bar{x}$ هو تقدير الامكان الاعظم الى P ؟

١١ ــ ٥ : لتكن $x_1, x_2, ..., x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع $N(0, \sigma^2)$. جد تقدير الامكان الاعظم الى σ^2 ثم جد تباين هذا التقدير

١١ ــ ٦ : جد تقدير الامكان الاعظم للمعلمة β في توزيع $G(\alpha, \beta)$ بفرض ان α معلومة ثم جد تباين هذا التقدير .

١١ ـ ٧ : اذا علمت ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من توزيع پواسون بالمعلمة θ . بين ان $\bar{x}$ هو التقدير الغير متحيز ذو اقل تباين الى θ .

١١ ـ ٨ : افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(0, \sigma^2)$. هل يمكن القول ان $\frac{1}{n} \Sigma x_i^2$ هو التقدير الغير متحيز ذو اقل تباين الى σ^2 ؟

١١ ـ ٩ : كانت معطيات عينة عشوائية ذات حجم 17 مفردة مسحوبة من $N(\mu, \sigma^2)$. هي $S^2 = 6.2, \bar{x} = 5.3$ جد حدي الثقة بمعامل ثقة 0.95 الى كل من σ^2, μ .

١١ ـ ١٠ : اذا علمت ان $\bar{x} = 18$ يمثل الوسط الحسابي لقياسات عينة عشوائية ذات حجم 20 مسحوبة من $N(\mu, 25)$. جد حدي الثقة بمعامل ثقة 99 % الى μ .

١١ ـ ١١ : افرض ان $\bar{x}$ يمثل الوسط الحسابي لقياسات عينة عشوائية ذات حجم n مسحوبة من $N(\mu, 25)$. جد اقرب عدد صحيح الى n بحيث ان $P_r(\bar{x} - 1 < \mu < \bar{x} + 1) = 0.99$.

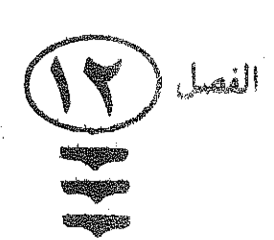
١١ ـ ١٢ : افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\mu_1, \sigma^2)$ وان $S_1^2, \bar{x}$ يمثلان على التوالي الوسط والتباين لهذه العينة . وافرض ان $y_1, y_2, \ldots, y_m$ تمثل قياسات عينة اخرى مسحوبة من $N(\mu_2, \sigma^2)$ وان $S_2^2, \bar{y}$ يمثلان على التوالي الوسط والتباين لهذه العينة . وبفرض ان المجتمعين مستقلان يطلب ايجاد فترة ثقة بمعامل ثقة $(1 - \alpha)$ للفرق بين متوسطي هذين المجتمعين $(\mu_1 - \mu_2)$.

١١ ـ ١٣ : اختيرت عينة عشوائية من المفردات ذات حجم 10 من $N(\mu, \sigma^2)$ وكانت نتائج القياس لمفردات العينة كما يلي :
10.7 ، 12.6 ، 9.3 ، 9.5 ، 11.3 ، 12.2 ، 11.5 ، 11.1 ، 10.4 ، 10.2 يطلب ايجاد فترة ثقة بمعامل ثقة 0.95 الى كل من σ^2, μ .

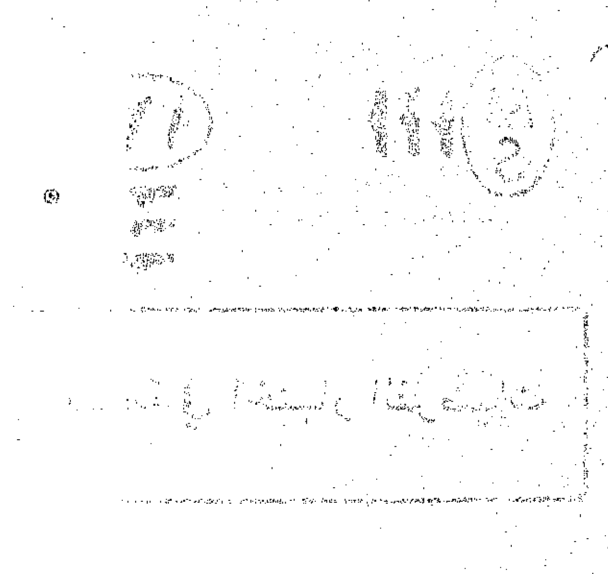
١١ ـ ١٤ : اذا علمت ان معطيات عينة ذات حجم 10 مسحوبة من $N(\mu_1, \sigma^2)$ كانت $S_1^2 = 8, \bar{x} = 5$ وان معطيات عينة اخرى ذات حجم 12 مسحوبة من $N(\mu_2, \sigma^2)$ كانت $S_2^2 = 9, \bar{y} = 6$. جد فترة ثقة بمعامل ثقة 0.95 الى $(\mu_1 - \mu_2)$.

١١ ـ ١٥ : افرض ان S_1^2 يمثل تباين عينة ذات حجم n مسحوبة من $N(\mu_1, \sigma_1^2)$ وان S_2^2 يمثل تباين عينة ذات حجم m مسحوبة من $N(\mu_2, \sigma_2^2)$. فاذا علمت ان μ_2, μ_1 مجهولين وان المجتمعين مستقلان يطلب ايجاد فترة ثقة بمعامل ثقة $(1 - \alpha)$ للنسبة σ_1^2 / σ_2^2 .





مقدمة في اختبار الفرضيات



الفصل الثاني عشر

مقدمة في اختبار الفرضيات

تطرقنا في الفصل السابق وبشكل موجز لنظرية التقدير مع عرض لاهم الصفات التي يجب ان يتمتع بها التقدير الجيد . كذلك تم عرض اهم الطرق التي من شأنها التوصل لذلك التقدير بالاضافة الى استعراض موجز لفترات الثقة واسلوب تكوينها . في هذا الفصل سوف نتطرق وبشكل موجز ايضاً لموضوع اختبار الفرضيات الذي يعتبر الشق الثاني المكمل لنظرية التقدير في موضوع الاستدلال الاحصائي .

١٢ ـ ١ : مفهوم اختبار الفرضيات The concept of testing hypotheses

افرض ان مصنعا لانتاج نوع من الصمامات الالكترونية يعتمد مواصفات قياسية في الانتاج مفادها ان متوسط عمر الصمام المنتج وفق هذه المواصفات هو 1500 ساعة اشتغال . اقترح احد مهندسي المصنع اجراء بعض التعديلات في تلك المواصفات مدعياً ان هذا الاجراء سوف يرفع من متوسط عمر الصمام الى قدر اكبر من 1500 ساعة اشتغال الامر الذي سيؤدي الى ارتفاع الطلب عليه وبالتالي زيادة ارباح هذا المصنع . من الناحية العملية يلاحظ ان اقتراح هذا المنهدس جيد بسبب ان ادعاءه سيؤدي الى رفع متوسط عمر اشتغال الصمام ، الا انه من ناحية اخرى لايمكن القبول وبشكل نهائي بمنطق من هذا النوع لاسباب قد تتعلق بعوامل الصدفة ، وليس لاسباب فنية (التعديلات في مواصفات الانتاج) ، ادت صدفة الى رفع متوسط عمر الصمام المنتج . عليه وبهدف التأكد من صحة ادعاء هذا المهندس لابد من اختبار هذا الادعاء .

يلاحظ من خلال هذا المثال اننا سوف نتعامل مع مجتمعين احصائيين الاول يمثل مجتمع الصمامات المنتجة وفق اسلوب الانتاج المعتمد والثاني يمثل مجتمع الصمامات المنتجة وفق اسلوب الانتاج المقترح من قبل المهندس .

وان الامر يتطلب الاجابة على السؤال التالي : هل ان متوسط عمر الصمام المنتج وفق اسلوب الانتاج المقترح هو اكبر من 1500 ساعة ام غير ذلك ؟ وهذا يعني ان هنالك فرضية يتطلب الامر اختبارها مفاد هذه الفرضية هو ان متوسط عمر الصمام وفق اسلوب الانتاج المقترح هو اكبر من 1500 ساعة . وبهدف اختبار هذه الفرضية يتوجب الامر سحب عينة عشوائية من الصمامات المنتجة وفق اسلوب الانتاج المقترح وعلى ضوء قياسات هذه العينة يتم اتخاذ القرار بشأن قبول او رفض ادعاء المهندس .

فعلى فرض ان قياسات اعمار الصمامات لعينة عشوائية مسحوبة من انتاج الاسلوب المقترح قوامها 25 صمام بينت ان $\bar{x} = 1600$ ساعة وان $S = 625$ ساعة وبفرض ان العينة قد اختيرت من مجتمع طبيعي فان فترة الثقة لمتوسط عمر الصمام المنتج وفق اسلوب الانتاج المقترح (المجتمع الثاني) بمعامل ثقة 95 % هي 1342 ساعة الى 1858 ساعة . ووفقاً لهذه المعطيات يمكن القول ان متوسط عمر الصمام المنتج وفق اسلوب الانتاج المقترح هو نتيجة لقياسات عينة مسحوبة من مجتمع طبيعي متوسطه 1500 ساعة حيث ان $\mu = 1500$ عدد معرف في الفترة (1342 , 1858) . وهذا يعني انه ليس لدينا اي سبب كاف لرفض الفرضية القائلة بان متوسط عمر الصمام هو 1500 ساعة والقبول بادعاء المهندس . ولمعطيات المثال السابق وبفرض ان $n = 25 , S = 125 , \bar{x} = 1600$ فان فترة الثقة لمتوسط عمر الصمام المنتج وفق اسلوب الانتاج المقترح هي (1548.4 , 1651.6) وهذا يعني ان هنالك دليل كافٍ للقول بان متوسط عمر الصمام المنتج وفق الاسلوب المقترح هو نتيجة لقياسات عينة مسحوبة من مجتمع طبيعي متوسطه اكبر من 1500 ساعة وبالتحديد قد يتراوح ضمن الفترة (1548.4 , 1651.6) مما يدعونا الامر الى رفض الفرضية القائلة بان $\mu = 1500$ والقبول باقتراح المهندس .

١٢ ـ ٢ : الفرضية الاحصائية Statistical hypothesis

بشكل عام يمكن القول ان الفرضية الاحصائية هي ادعاء او تصريح يتعلق بالتوزيع الاحتمالي (المجتمع الاحصائي) لمتغير عشوائي . فمثلاً الادعاء القائل بان « المتغير العشوائي X يتوزع كتوزيع طبيعي بوسط μ وتباين σ^2) هو فرضية احصائية . الادعاء القائل بان « متوسط قيم متغير عشوائي يتوزع كتوزيع

بواسون هو $\lambda = 5$ » هو فرضية احصائية . والادعاء القائل بان « متوسط عمر الصمام المنتج وفق مواصفات قياسية هو 1500ساعة» هو فرضية احصائية. وهذا يعني ان الفرضية الاحصائية هي ادعاء يخص تلك المعالم التي تصف التوزيع الاحتمالي لمتغير عشوائي والتي نرغب في التحري عنها على اساس المعلومات المستقاة من عينة عشوائية مسحوبة من ذلك التوزيع . وبشكل عام وكما هو متداول في اغلب الادبيات الاحصائية فانه يرمز لاية فرضية احصائية بالرمز H . فللأدعاء الاول يمكن صياغة H بالشكل $H : X \sim N(\mu, \sigma^2)$ ، وللثاني فان $H : \lambda = 5$ ، وللثالث فان $H : \mu = 1500$.

والفرضية الاحصائية على نوعين فرضية بسيطة Simple hypothesis وفرضية مركبة Composite hypothesis . ويقال ان H فرضية بسيطة اذا وصفت هذه الفرضية التوزيع الاحتمالي بشكل متكامل وبخلاف ذلك يقال ان H فرضية مركبة . فمثلاً وعند اختبار فرضية تتعلق بمجتمع توزيعه $N(\mu, \sigma^2)$ فان الفرضية $H : \mu = \mu_0 , \sigma^2 = \sigma_0^2$ ، حيث ان σ_0^2 , μ_0 قيمتان معلومتان ، هي فرضية بسيطة كونها وصفت التوزيع الاحتمالي بشكل متكامل (اي انها وصفت معلمتي التوزيع وبنقطة واحدة لكل منهما هي μ_0 الى μ و σ_0^2 الى σ^2) في حين ان الفرضيات الثلاث التالية هي فرضيات مركبة : $H : \mu = \mu_0$ ، $H : \mu = \mu_0 , \sigma^2 > \sigma_0^2$ ، $H : \mu > \mu_0, \sigma^2 < \sigma_0^2$ بسبب ان الفرضية الاولى وصفت المعلمة μ بنقطة ولم تصف σ^2 ، والثانية وصفت المعلمة μ بنقطة الا انها وصفت المعلمة σ^2 بمجال ، والثالثة وصفت معلمتي التوزيع بمجال وليس بنقطة . كذلك فان الفرضية التي تصف التوزيع بشكل متكامل عدا وصفها لعدد آخر من معالم ذلك التوزيع ، ليكن V مثلاً ، يقال عنها فرضية مركبة بـ V درجة حرية . والفرضية من هذا النوع يتم النظر اليها على انها تجميع لفرضيات بسيطة او مركبة من عدة فرضيات بسيطة . فمثلاً الفرضية القائلة بان التوزيع الاحتمالي لمتغير عشوائي هو توزيع طبيعي بوسط $\mu = \mu_0$ هي فرضية مركبة من عدة فرضيات بسيطة كل منها تختلف عن الاخرى باختلاف قيمة σ^2 وفي هذه الحالة يقال ان H فرضية مركبة بدرجة حرية واحدة لانها وصفت μ ولم تصف σ^2 .

١٢ ـ ٢ ـ ١ : اختبار الفرضية الاحصائية Test of statistical hypothesis

ان اختبار اية فرضية مثل H عبارة عن طريقة او قاعدة لاتخاذ القرار بشأن H. فاذا كان $f(x, \theta)$ يمثل التوزيع الاحتمالي لمتغير عشوائي X وان $H : \theta = \theta_0$ تمثل ادعاء بشأن θ وان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من هذا التوزيع . عندئذٍ وعلى اساس قياسات هذه العينة فان اختبار H يمثل قاعدة لاتخاذ القرار بشأن H هذا القرار هو ذو حدين اما قبول H او رفضها .

١٢ ـ ٢ ـ ٢ : فرضية العدم والفرضية البديلة Null and alternative hypothesis

من خلال المثال الموضح في الفقرة (١٢ ـ ١) لاحظنا ان المسألة تطلبت التعامل مع مجتمعين احصائيين الاول تمثل بالصمامات الالكترونية المنتجة وفق الاسلوب المعتمد والثاني تمثل بالصمامات الالكترونية المنتجة وفق الاسلوب المقترح ، وان المطلوب هنا اختبار ادعاء هذا المهندس . في هذه الحالة نلاحظ اننا نقف امام ثلاث فرضيات ممكنة هي :

(١) : ان الاسلوب المقترح هو افضل من الاسلوب المعتمد .
(٢) : ان الاسلوب المعتمد هو افضل من الاسلوب المقترح .
(٣) : انه لايوجد فرق بين الاسلوبين .

ومن خلال عملية التقصي للحالات الثلاث نلاحظ ان الفرضيتان (١) ، (٢) متحيزتين ، فالاولى متحيزة لأدعاء المهندس والثانية متحيزة لمواصفات المصنع وفي هذه الحالة سوف نقف امام مسألة تحديد اي من الفرضيتين يتوجب اختبارها ، لكن نلاحظ ان الفرضية الثالثة ليست متحيزة لاي من الاسلوبين فهي تنص بعدم وجود فرق بين اسلوبي الانتاج لذا يمكن اعتمادها كفرضية اختبار بسبب كونها فرضية محايدة . ان الفرضية المحايدة وفق المفهوم اعلاه تسمى فرضية العدم وهي تمثل نقطة الانطلاق بالنسبة للقائم بعملية الاختبار قبل البدء بجمع البيانات اللازمة للاختبار . وغالباً مايرمز لفرضية العدم بالرمز H_0 . فاذا رمزنا لمتوسط عمر

الصمام المنتج وفق الاسلوب المعتمد بـ μ_0 ولمتوسط عمر الصمام المنتج وفق الاسلوب المقترح بـ μ عندئذٍ يمكن صياغة فرضية العدم بالشكل $H_0 : \mu = \mu_0$ او $H_0 : \mu - \mu_0 = 0$ وبذلك فان H_0 تمثل تلك الفرضية التي يتم اختبار امكانية رفضها بفرض انها صحيحة .

ان مسألة القبول بالفرضية H_0 تعني ضمناً رفض بديل او بدائل معينة لها وان مسألة رفض H_0 تعني ضمناً القبول ببديل او بدائل معينة لها وهذا يعني ان قبول او رفض H_0 يمثل نتيجة لاتخاذ قرار بين H_0 وبديل آخر لها . ان البديل الآخر الى H_0 يسمى بالفرضية البديلة وغالباً ما يرمز لها بالرمز H_1

وبشكل عام يمكن القول بان لاي اختبار احصائي توجد فرضيتان هما H_0 و H_1 وان الاهتمام يكون منصب على اختبار الفرضية H_0 ضد الفرضية H_1 . عليه ومن الناحية العملية يتوجب ملاحظة من هي H_0 ومن هي H_1 على ضوء المشكلة تحت الدراسة . فللمثال السابق فان $H_0 : \mu = \mu_0$ يمكن ان تختبر ضد احد البدائل التالية : $H_1 : \mu > \mu_0$ او $H_1 : \mu < \mu_0$ او $H_1 : \mu \neq \mu_0$. وعلى ضوء الفرضية البديلة فانه يقال ان الاختبار من جانب واحد one - tailed test اذا كانت H_1 احد البديلين الاول او الثاني السابقين ، ويقال ان الاختبار من جانبين two - tailed test اذا كانت H_1 تمثل البديل الثالث . ويراعى قدر الامكان وعند صياغة H_0 ان تكون فرضية بسيطة والابتعاد قدر الامكان عن صياغتها بشكل فرضية مركبة وحسب المشكلة تحت الدراسة .

١٢ ـ ٢ ـ ٣ : الخطأ من النوع الاول والثاني Two types of error

لتكن $x_1, x_2, \ldots, x_n$ قياسات عينة عشوائية من المفردات مسحوبة من مجتمع بدالة كثافة احتمالية $f(x, \theta)$ ، وافرض اننا نرغب في اختبار الفرضية $H_0 : \theta = \theta_0$ ضد $H_1 : \theta \neq \theta_0$. ان اتخاذنا لقرار بشأن قبول او رفض H_0 فان ذلك القرار قد استند في الحقيقة على معطيات تلك العينة . ان المؤشرات المستقاة على ضوء قياسات العينة فيما يخص المعلمة θ قد لاتكون صائبة دائماً الامر الذي قد يؤدي الى وقوع الباحث في نوعين رئيسين من الاخطاء هما الخطأ من النوع الاول والثاني .

ويعرف الخطأ من النوع الاول بأنه الخطأ الحاصل بسبب رفض H_0 عندما تكون صحيحة . في حين يعرف الخطأ من النوع الثاني بأنه الخطأ الحاصل بسبب قبول H_0 عندما تكون خاطئة ، وكما هو موضح في الجدول التالي :

القرار / الفرضية	رفض H_0	قبول H_0
H_0 صحيحة	قرار خاطيء (الخطأ من النوع الاول)	قرار صحيح
H_0 خاطئة	قرار صحيح	قرار خاطيء (الخطأ من النوع الثاني)

وغالباً ما يرمز لاحتمال حدوث الخطأ من النوع الاول بالرمز α ، ولاحتمال حدوث الخطأ من النوع الثاني بالرمز β . وهذا يعني ان :

α = احتمال حدوث الخطأ من النوع الاول

= احتمال رفض H_0 عندما تكون صحيحة .

β = احتمال حدوث الخطأ من النوع الثاني

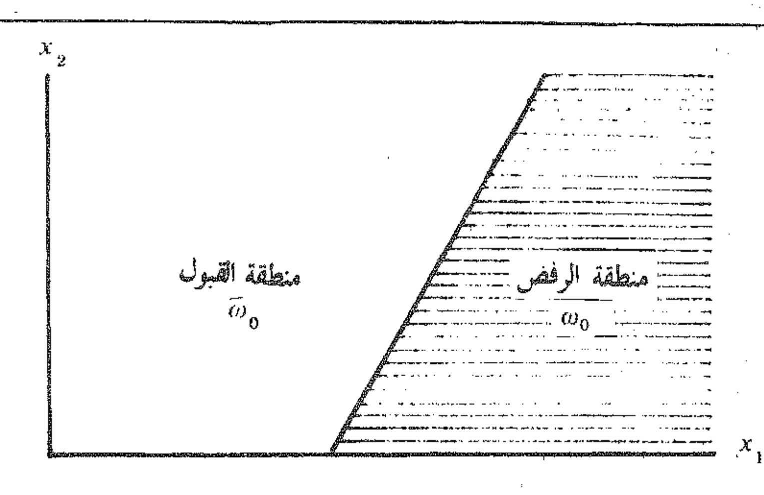
∴ = احتمال قبول H_0 عندما تكون خاطئة .

١٢ ـ ٢ ـ ٤ : المنطقة الحرجة Critical region

كما سبق ذكره في الفقرة (١٢ ـ ٢ ـ ١) فان اختبار اية فرضية مثل H_0 يمثل طريقة او قاعدة لاتخاذ القرار بشأن قبول او رفض H_0 . في الحقيقة فان هذه القاعدة تتمثل في تحديد منطقة ، لتكن W_0 مثلاً ، في فضاء عينة اقليدي ذو n بُعد ، ليكن S مثلاً ، بحيث لو وقعت نقطة العينة* Sample point في المنطقة W_0 لادى ذلك الى رفض H_0 وبالعكس تقبل H_0 اذا وقعت نقطة العينة خارج

* اذا كانت $x_1 , x_2 , \ldots , x_n$ تمثل قياسات عينة عشوائية مسحوبة من $f(x , \theta)$ عندئذٍ يمكن تمثيل هذه القياسات بنقطة في فضاء اقليدي ذو n بُعد . هذه النقطة تسمى نقطة العينة وهي في الحقيقة تمثل مؤشر احصائي يعتمد على قياسات العينة فقط مثل $\Sigma x_i^2 , \Sigma x_i , \bar{x}$ وغيرها .

فاذا امكن لنا تجزئة فضاء العينة $\hat{S}$ الى مجموعتين غير مشتركتين مثل W_0 و $\bar{W}_0 = S - W_0$، وان $W_0 \cap \bar{W}_0 = \phi$ ، عندئذٍ ترفض H_0 اذا وقعت نقطه العينة في W_0 وتقبل H_0 اذا وقعت نقطة العينة في $\bar{W}_0$. ان اية منطقة تتميز بما سبق ذكره تسمى منطقة حرجة او منطقة رفض H_0 . فمثلاً اذا كانت x_1, x_2 تمثل قياستين فان المنطقة الحرجة يمكن تمثيلها بالشكل (١٢ ــ ١) :



شكل (١٢ ــ ١) : توضيح لمنطقة حرجة

وغالباً مايكون الاهتمام منصب في تنفيذ اختبار معين اي اختيار منطقة حرجة بالشكل الذي يقلل من فرص الوقوع في الخطأ من النوع الاول والثاني معاً . الا ان ذلك امر صعب التنفيذ من الناحية العملية . فمسألة تقليل فرص الوقوع في احدهما قد يؤدي الى زيادة فرص الوقوع في الخطأ من النوع الاخر . عليه وبفرض ان قرارنا هو قبول H_0 دائماً فان ذلك يمكننا من تقليص الخطأ من النوع الاول وجعله صفراً الا ان ذلك قد يؤدي الى جعل احتمال حدوث الخطأ من النوع الثاني β عند قيمته الكبرى . عليه ومن الناحية العملية فانه غالباً مايصار الى تثبيت احتمال الخطأ من النوع الاول α عند قيمة معينة محددة سلفاً ثم يتم اختيار تلك المنطقة الحرجة من بين مجموعة مناطق حرجة متاحة (كل منها باحتمال خطأ من النوع الاول هو α) التي فيها احتمال الخطأ من النوع الثاني اقل مايمكن .

١٢ ـ ٢ ـ ٥ : مستوى المعنوية level of significance

ان احتمال الوقوع في الخطأ من النوع الاول (α = احتمال رفض H_0 عندما تكون صحيحة) يسمى مستوى المعنوية للاختبار ، او حجم المنطقة الحرجة ، او حجم الاختبار . وهذا يعني ان مستوى المعنوية يمثل قيمة α وغالباً مايتم تحديد قيمة α سلفاً وقبل البدء بتنفيذ الاختبار كان تكون مساوية الى 1% او 5% او غير ذلك . وكلما كانت قيمة α صغيرة فذلك معناه تقليص في حجم المنطقة الحرجة اي زيادة حجم منطقة قبول H_0 اي مانعنيه تقليل فرص الوقوع في الخطأ من النوع الاول . فاذا رمزنا لنقطة العينة بالرمز X حيث $X = (x_1, x_2, \ldots, x_n)$ وان L_0 تمثل دالة الامكان لقياسات العينة تحت فرضية العدم H_0 فان :

$$\alpha = P_r(X \in W_0 \mid H_0) = \int_{W_0} L_0 dx$$

حيث ان

$$\int dx \text{ تعني } \int\int \ldots \int dx_1 dx_2 \ldots dx_n$$

١٢ ـ ٢ ـ ٦ : قوة الاختبار power of the test

تعرف قوة الاختبار (قوة المنطقة الحرجة) بانها احتمال وقوع نقطة العينة في المنطقة الحرجة عندما تكون H_0 خاطئة وان اية فرضية اخرى بديلة مثل H_1 تكون صحيحة . ويتضح من هذا التعريف ان قوة الاختبار تعتمد على مدى صحة الفرضية H_1 . عليه فان قوة الاختبار لاختبار H_0 ضد H_1 هي :

P. o. t = احتمال (رفض H_0 عندما H_1 صحيحة)

= احتمال (رفض H_0 عندما H_0 خاطئة)

= 1 − احتمال (قبول H_0 عندما H_0 خاطئة)

= 1 − احتمال (حدوث الخطأ من النوع الثاني)

$= 1 - \beta$

مما تقدم يتضح ان قوة الاختبار تمثل معيار لشدة رفضنا للفرضية H_0 عندما تكون H_0 خاطئة . كذلك فان قوة الاختبار تزداد من خلال تقليصنا لفرص الوقوع في الخطأ من النوع الثاني اي تقليل احتمال حدوث الخطأ من النوع الثاني β .

وبالرموز وبفرض ان L_1 تمثل دالة الامكان لقياسات العينة تحت الفرضية H_1 فان :

$$\beta = P_r(X \in W_0 \mid H_1) = \int_{\bar{W}_0} L_1 \, dx$$

لكن

$$\int_{W_0} L_1 dx + \int_{\bar{W}_0} L_1 dx = P_r(S) = 1$$

عندئذٍ فان

$$\int_{W_0} L_1 dx = 1 - \int_{\bar{W}_0} L_1 dx = 1 - \beta$$

اي ان

$$P_r(X \in W_0 \mid H_1) = 1 - \beta$$

مثال (١) : افرض ان X متغير عشوائي يتوزع كتوزيع اسي بالمعلمة θ وافرض ان المنطقة الحرجة لاختبار الفرضية $H_0 : \theta = 2$ ضد الفرضية $H_1 : \theta = 1$ على اساس قياسة واحدة مسحوبة من هذا المجتمع هي $x \geq 1$. جد قيمة α, β وقوة الاختبار .

الحل :

واضح من معطيات هذا المثال ان :

$$W_0 = \{x : x \geq 1\} \quad \therefore \bar{W}_0 = \{x : x < 1\}$$

وان

$$f(x, \theta) = \theta e^{-\theta x} \quad \therefore L_0 = 2e^{-2x}, L_1 = e^{-x}$$

$$\therefore \alpha = P_r(x \in W_0 \mid H_0) = P_r(x \geq 1 \mid \theta = 2)$$

$$\therefore \alpha = \int_1^\infty 2e^{-2x} dx = 0{\cdot}1353352$$

كذلك فان

$$\beta = P_r(x \in \bar{W}_0 \mid H_1) = P_r(x < 1 \mid \theta = 1)$$

$$\therefore \beta = \int_0^1 e^{-x} dx = 0{\cdot}6321206 \rightarrow 1 - \beta = 0{\cdot}3678794$$

مثال (٢) : لمعطيات المثال (١) وبفرض ان المنطقة الحرجة لاختبار H_0 هي $x \geq x_0$. جد x_0 بحيث ان $\alpha = 5\%$ ثم جد β وقوة الاختبار .

الحل :

$$\alpha = \int_{x_0}^{\infty} 2e^{-2x}\,dx = 0{\cdot}05$$

$$\Longrightarrow -[e^{-2x}]_{x_0}^{\infty} = 0{\cdot}05$$

$$\Longrightarrow -[0 - e^{-2x_0}] = 0{\cdot}05 \Longrightarrow e^{-2x_0} = 0{\cdot}05$$

$$\therefore -2x_0 = \ln(0{\cdot}05) = -2{\cdot}9957323$$

$$\therefore x_0 = 1{\cdot}5$$

$$\therefore W_0 = \{x : x \geq 1{\cdot}5\},\ \bar{W}_0 = \{x : x < 1{\cdot}5\}$$

$$\therefore \beta = \int_{0}^{1\cdot5} e^{-x}dx = 0{\cdot}2231301$$

$$\therefore 1 - \beta = 0{\cdot}7768699$$

مثال (٣) : افرض ان X متغير عشوائي يتوزع كتوزيع پواسون بالمعلمة λ . وافرض ان المنطقة الحرجة لاختبار الفرضية $H_0 : \lambda = 1$ ضد $H_1 : \lambda = 2$ على اساس قياسة واحدة مسحوبة من هذا المجتمع هي $x > 2$. جد قيمة α, β وقوة الاختبار .

الحل :

واضح من معطيات هذا المثال ان :

$$W_0 = \{x : x > 2\},\ \bar{W}_0 = \{x : x \leq 2\}$$

وان

$$L_0 = P(x, \lambda \mid H_0) = \frac{e^{-1}}{x!}, \quad L_1 = P(x, \lambda \mid H_1) = \frac{2^x e^{-2}}{x!}$$

$$\therefore \alpha = P_r(x \in W_0 \mid H_0) = P_r(x > 2 \mid \lambda = 1)$$

$$= 1 - P_r(x \leq 2 \mid \lambda = 1) = 1 - \sum_{x=0}^{2} \frac{e^{-1}}{x!} = 0{\cdot}0803015$$

وان

$$\beta = P_r(x \in w_0 \mid H_1) = P_r(x \leq 2 \mid \lambda = 2)$$

$$= \sum_{x=0}^{2} P(x \mid \lambda = 2) = \sum_{x=0}^{2} \frac{2^x e^{-2}}{x!} = 0{\cdot}676676$$

فاذن

$$1 - \beta = 0{\cdot}323324$$

١٢ ـ ٣ : الاختبارات المثلى Optimum tests

ذكرنا في الفقرة (١٢ ـ ٢ ـ ٦) ان قوة الاختبار تمثل معيار لشدة رفض H_0 عندما تكون خاطئة . وغالباً مايكون الهدف منصب في البحث عن ذلك الاختبار الذي تكون فيه β اقل مايمكن بفرض ثبات α . او بشكل مكافيء ان $(1 - \beta)$، اي قوة الاختبار ، تكون اكبر مايمكن . ان الاختبار الذي يحقق ماتقدم يسمى افضل اختبار best test .

١٢ ـ ٣ ـ ١ : الاختبار الاكثر قوة Most powerful test (MPT)

افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع بدالة كثافة احتمالية $f(x, \theta)$. وافرض اننا نرغب في اختبار الفرضية البسيطة $H_0: \theta = \theta_0$ ضد الفرضية البسيطة $H_1: \theta = \theta_1$. عندئذٍ يقال ان اختبار الفرضية H_0 ضد H_1 هو الاختبار الاكثر قوة ذات حجم α اذا كان :

$$P_r(x \in \omega_0 \mid H_0) = \int_{\omega_0} L_0 \, dX = \alpha$$

وان

$$P_r(x\in\omega_0 \mid H_1) \geq P_r(x\in w \mid H_1)$$

حيث ان w تعني اي منطقة حرجة اخرى بحيث ان ،

$$P_r(X\in\omega \mid H_0) = \int_{\omega} L_0\, dX = \alpha$$

وعندئذٍ يقال ان المنطقة الحرجة w_0 هي المنطقة الحرجة الاكثر قوة من اية منطقة حرجة اخرى مثل w_0 بفرض ثبات حجم الاختبار α في كافة الاحوال . وببساطة فان الاختبار الاكثر قوة هو ذلك الاختبار الذي تكون فيه قوة الاختبار $(1-\beta)$ اكبر من قوة الاختبار لأي اختبار آخر متاح بفرض ثبات حجم الاختبار α في كلا الحالتين .

١٢ ـ ٢ ـ ٢ : الاختبار الاكثر قوة بانتظام Uniformly most powerful test (U. M. P. T)

افرض اننا نرغب في اختبار الفرضية البسيطة $H_0: \theta = \theta_0$ ضد الفرضية المركبة البديلة $H_1: \theta \neq \theta_0$. عندئذٍ يقال ان اختبار الفرضية H_0 ضد H_1 هو الاختبار الاكثر قوة بانتظام ذات حجم α اذا كان :

$$P_r(X\in\omega_0 \mid H_0) = \int_{\omega_0} L_0\, dX = \alpha$$

وان

$$P_r(X\in\omega_0 \mid H_1) \geq P_r(X\in\omega / H_1) \quad \forall\, \theta \neq \theta_0$$

حيث ان w تعني اي منطقة حرجة اخرى بحيث ان

$$P_r(X\in\omega \mid H_0) = \int_{\omega} L_0\, dX = \alpha \quad \forall\, \theta \neq \theta_0$$

١٢ ـ ٤ : مبرهنة نيمان ـ بيرسون Neyman-pearson lemma

افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع بدالة كثافة احتمالية $f(x, \theta)$ (او بدالة كتلة احتمالية مثل $P(x, \theta)$) . وافرض اننا نرغب في اختبار الفرضية البسيطة $H_0 : \theta = \theta_0$ ضد الفرضية البسيطة $H_1 : \theta = \theta_1$. ليكن K عدد موجب وان ω_0 تمثل منطقة حرجة ذات حجم α بحيث ان :

$$\omega_0 = \left\{ X \in S : \frac{f(x, \theta_1)}{f(x, \theta_0)} \geq K \right\}$$

$$= \left\{ X \in S : \frac{L_1}{L_0} \geq K \right\} \qquad \ldots (I)$$

وان

$$\bar{\omega}_0 = \left\{ X \in S : \frac{L_1}{L_0} \leq K \right\} \qquad \ldots (II)$$

حيث ان L_1, L_0 تمثلان دالتي الامكان لقياسات العينة $X = (x_1, x_2, \ldots, x_n)$ تحت H_0 ، H_1 على التوالي . عندئذٍ يقال ان المنطقة الحرجة ω_0 بقوة $(1 - \beta)$ هي المنطقة الحرجة الاكثر قوة لاختبار H_0 ضد H_1 .

البرهان :

لتكن ω منطقة حرجة اخرى ذات حجم $\alpha^* \leq \alpha$ بقوة $(1 - \beta^*)$. ان المطلوب برهنته هنا ان $(1 - \beta) \geq (1 - \beta^*)$ اي ان ω_0 هي المنطقة الحرجة الاكثر قوة من ω .

كما هو معلوم فان

$$\alpha = P_r(X \in \omega_0 \mid H_0) = \int_{\omega_0} L_0 \, dX$$

وان قوة المنطقة الحرجة ω_0 هي

$$1 - \beta = P_r (X \in \omega_0 \mid H_1) = \int_{\omega_0} L_1 \, dX$$

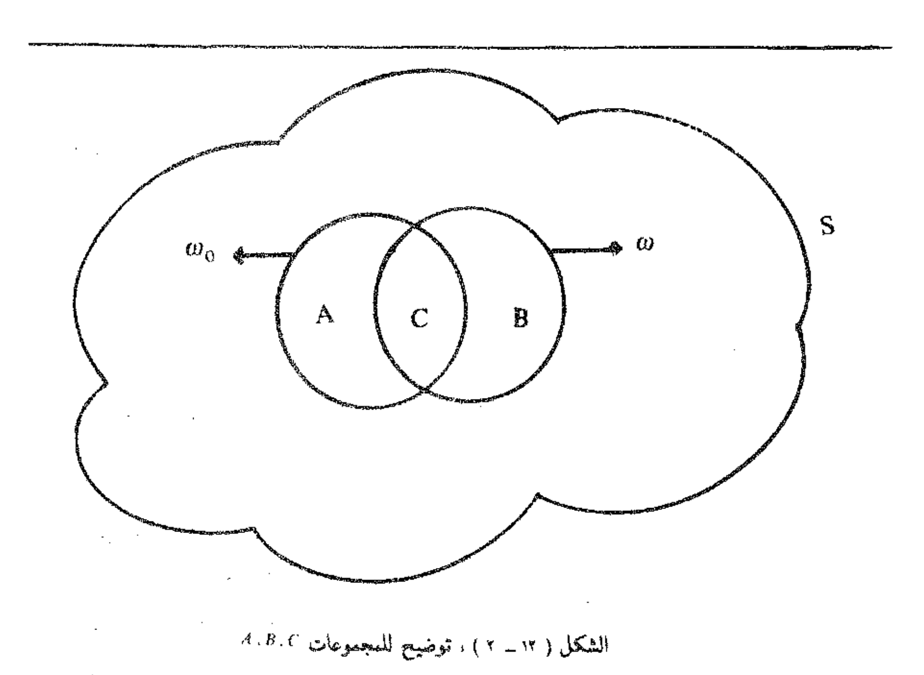
كذلك فان

$$\alpha^* = P_r (X \in \omega \mid H_0) = \int_{\omega} L_0 \, dX$$

وان

$$1 - \beta^* = P_r (X \in \omega \mid H_1) = \int_{w} L_1 \, dX$$

الان تأمل المجموعات الثلاث التالية المعرفة في S ولتكن C, B, A بحيث ان $A \cap B = C$ وكما موضح في الشكل (١٢ ـ ٢) .



الشكل (١٢ ـ ٢) ، توضيح للمجموعات A, B, C

الان بفرض ان $\omega_0 = A \cup C$ وان $\omega = B \cup C$ فاذا كانت $\alpha^* \le \alpha$ فذلك يعني ان :

او ان

$$\int_{w} L_0\, dX \le \int_{\omega_0} L_0\, dX \rightarrow \int_{B \cup C} L_0\, dX \le \int_{A \cup C} L_0\, dX$$

$$\int_{B} L_0\, dX \le \int_{A} L_0\, dX \rightarrow \int_{A} L_0\, dX \ge \int_{B} L_0\, dX$$

الان من المجموعة (I) فان $\omega_0 = \{ X \in S : L_1 \ge L_0 K \}$ طالما ان
وان $L_0 > 0$

$$\int_{A} L_1\, dX \ge \int_{A} K L_0\, dX \rightarrow \int_{A} L_1\, dX \ge K \int_{A} L_0\, dX$$

وهذا يعني ان

$$\int_{A} L_1\, dX \ge K \int_{A} L_0\, dX \ge K \int_{B} L_0\, dX$$

كذلك من المجموعة (II) فان $\bar{\omega}_0 = \{ X \in S : L_1 \le K L_0)$
وان

$$\int_{\bar{\omega}_0} L_1\, dX \le K \int_{\bar{\omega}_0} L_0\, dX$$

فاذن

$$\int_{B} L_1\, dX \le K \int_{B} L_0\, dX \le \int_{A} L_1\, dX$$

اي ان

$$\int_{B} L_1\, dX \le \int_{A} L_1\, dX$$

وباضافة المقدار $\int_{C} L_1\, dX$ لطرفي المتباينة الاخيرة نحصل على :

$$\int_{B} L_1\, dX + \int_{C} L_1\, dX \le \int_{A} L_1\, dX + \int_{C} L_1\, dX$$

$$\rightarrow \int_{B \cup C} L_1\, dX \le \int_{A \cup C} L_1\, dX \rightarrow \int_{w} L_1\, dX \le \int_{\omega_0} L_1\, dX$$

اي ان

$$1 - \beta^* \le 1 - \beta \rightarrow \beta^* \ge \beta$$

وهذا يعني ان المنطقة الحرجة ω_0 اي $\frac{L_1}{L_0} \geq K$ هي افضل منطقة حرجة ذات حجم α لاختبار H_0 .

وعلى ضوء هذه المبرهنة فانه يصار الى رفض H_0 اذا كانت $\frac{L_1}{L_0} \geq K$ والقبول بالفرضية H_1 . وفيما يلي بعض الامثلة التي توضح ذلك .

مثال (٤) : افرض ان $x_1 , x_2 , \ldots , x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع اسي بالمعلمة θ . جد المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار الفرضية $H_0 : \theta = \theta_0$ ضد الفرضية $H_1 : \theta = \theta_1 < \theta_0$ ثم جد قوة الاختبار .

الحل : ان

$$f(x , \theta) = \theta e^{-\theta x} , x \geq 0$$

وان

$$L_0 = \prod_{i=1}^{n} f(x_i , \theta = \theta_0) = \prod_{i=1}^{n} \theta_0 e^{-\theta_0 x_i} = \theta_0^n . e^{-\theta_0 \Sigma x_i}$$

$$L_1 = \prod_{i=1}^{n} f(x_i , \theta = \theta_1) = \prod_{i=1}^{n} \theta_1 e^{-\theta_1 x_i} = \theta_1^n . e^{-\theta_1 \Sigma x_i}$$

وحسب مبرهنة نيمان ـ بيرسون فان افضل منطقة حرجة لاختبار H_0 هي $\frac{L_1}{L_0} \geq K$ اي ان :

$$\frac{\theta_1^n . e^{-\theta_1 \Sigma x_i}}{\theta_0^n . e^{-\theta_0 \Sigma x_i}} \geq K \rightarrow \left(\frac{\theta_1}{\theta_0} \right)^n . e^{(\theta_0 - \theta_1) \Sigma x_i} \geq K$$

او ان

$$e^{(\theta_0 - \theta_1) \Sigma x_i} \geq \left(\frac{\theta_0}{\theta_1} \right)^n . K$$

$$\therefore (\theta_0 - \theta_1) \Sigma x_i \geq n (\ln \theta_0 - \ln \theta_1) + \ln K = K^*$$

وحيث ان $\theta_0 - \theta_1 > 0$ فاذن

$$\Sigma x_i \geq \frac{K^*}{\theta_0 - \theta_1} = \lambda$$

وهذا يعني ان افضل منطقة حرجة لاختبار H_0 هي $\Sigma x_i \geq \lambda$ اي يتم رفض H_0 اذا كان مجموع قياسات العينة اكبر من او يساوي العدد λ . الهدف الان هو ايجاد λ بحيث ان λ تحقق :

$$\alpha = P_r(X \in \omega_0 \mid H_0) = \int_{\omega_0} L_0 \, dX$$

ويتضح من هذا المثال ان نقطة العينة X تتمثل بالمؤشر الاحصائي Σx_i . عليه فان :

$$P_r(\Sigma x_i \geq \lambda \mid H_0 : \theta = \theta_0) = \alpha$$

او ان

$$P_r(2\theta \Sigma x_i \geq 2\theta \lambda \mid H_0 : \theta = \theta_0) = \alpha$$

ولغرض حساب هذا الاحتمال فان الامر يتوجب اولاً تحديد التوزيع الاحتمالي الى $2\theta \Sigma x_i$ وسوف نستخدم اسلوب الدالة المولدة للعزوم وكما يلي : افرض ان $Y = \Sigma x_i$ فاذن

$$M_Y(t) = Ee^{tY} = Ee^{t \Sigma x_i} = \prod_{i=1}^{n} M_{X_i}(t)$$

لكن

$$M_{X_i}(t) = \frac{\theta}{\theta - t} \quad \therefore M_Y(t) = \prod_{i=1}^{n} \left(\frac{\theta}{\theta - t} \right) = \left(\frac{\theta}{\theta - t} \right)^n$$

او ان

$$M_Y(t) = \left(1 - \frac{1}{\theta} t \right)^{-n}$$

عليه وحسب خصائص الدالة المولدة للعزوم فان :

$$M_{2\theta Y}(t) = M_Y(2\theta t)$$

$$= \left(1 - \frac{1}{\theta}(2\theta t)\right)^{-n} = (1 - 2t)^{-n}$$

والصيغة الاخيرة تمثل الدالة المولدة لعزوم توزيع مربع كاي بـ (2 n) درجة حرية . وذلك يعني ان :

$$2\,\theta\,\Sigma\,x_i \sim \chi^2_{(2n)}$$

عليه فان :

$$\alpha = P_r(2\,\theta\,\Sigma\,x_i \geq 2\,\theta\,\lambda \mid H_0)$$

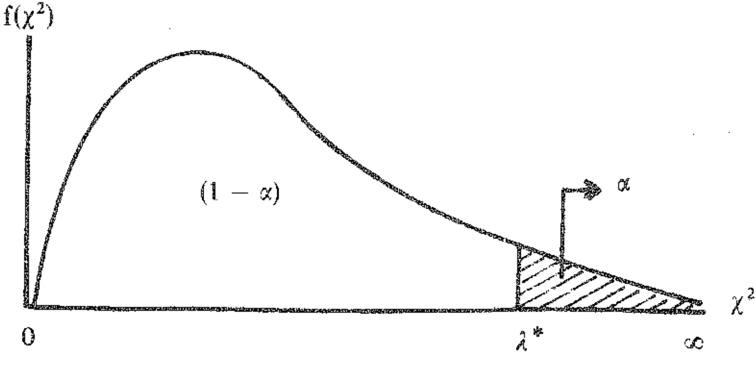
$$\therefore \alpha = P_r(\chi^2_{(2n)} \geq \lambda^* \mid H_0),\ \lambda^* = 2\,\theta_0\,\lambda$$

$$= 1 - P_r(\chi^2_{(2n)} \leq \lambda^* \mid H_0) \rightarrow 1 - \alpha = P_r(\chi^2_{(2n)} \leq \lambda^* \mid H_0)$$

ويتضح مما تقدم ان $P_r(\chi^2_{(2n)} \leq \lambda^* \mid H_0)$ تمثل قيمة التراكم الاحتمالي في توزيع مربع كاي معرف بـ (2n) درجة حرية لغاية λ^* بحيث ان قيمة هذا التراكم هي (α − 1) وهذا يعني ان λ^* تمثل قيمة من قيم χ^2 النظرية (حسب جدول هذا التوزيع) اي ان :

$$\lambda^* = \chi^2_{(2n)}(\alpha)$$

وكما هو موضح في الشكل (١٢ ــ ٣) :



الشكل (١٢ ــ ٣) : توضيح لقيم λ^* .

عليه فان

$$2\,\theta_0\,\Sigma\,x_i \geq 2\theta_0\,\lambda = \lambda^* = \chi^2_{(2n)}(\alpha)$$

وبذلك فان المنطقة الحرجة ستكون :

$$\omega_0 = \left\{ X : \Sigma\,x_i \geq \frac{\chi^2_{(2n)}(\alpha)}{2\theta_0} \right\}$$

وهذا يعني ان H_0 ترفض اذا كان مجموع القياسات اكبر من او يساوي $\frac{\chi^2_{(2n)}(\alpha)}{2\theta_0}$.

وبهدف حساب قوة الاختبار فان ذلك يتم وفق الآتي :

$$\begin{aligned}
\beta &= P_r(X \in \omega_0 \mid H_1 : \theta = \theta_1) \\
&= P_r(\Sigma\,x_i \leq \lambda \mid H_1 : \theta = \theta_1) \\
&= P_r(2\,\theta\,\Sigma\,x_i \leq 2\,\theta\,\lambda \mid H_1 : \theta = \theta_1) \\
&= P_r(2\,\theta_1\,\Sigma\,x_i \leq 2\,\theta_1\,\lambda)
\end{aligned}$$

لكن

$$\lambda^* = \chi^2_{(2n)}(\alpha) = 2\,\theta_0\,\lambda$$

$$\therefore \lambda = \frac{\chi^2_{(2n)}(\alpha)}{2\theta_0}$$

فاذن

$$\beta = P_r\left(\chi^2_{(2n)} \leq 2\theta_1 \cdot \frac{\chi^2_{(2n)}(\alpha)}{2\theta_0}\right)$$

$$= P_r\left(\chi^2_{(2n)} \leq \left(\frac{\theta_1}{\theta_0}\right) \cdot \chi^2_{(2n)}(\alpha)\right)$$

وبذلك فان قوة الاختبار هي :

$$1 - \beta = 1 - P_r\left(\chi^2_{(2n)} \leq \left(\frac{\theta_1}{\theta_0}\right) \cdot \chi^2_{(2n)}(\alpha)\right)$$

مثال (٥) لمعيطات المثال (٤) وبفرض ان $\theta_1 = 2, \theta_0 = 3, n = 4$ ، $\alpha = 0.05$. جد المنطقة الحرجة لاختبار H_0 ثم احسب قوة الاختبار

الحل :

من جداول توزيع مربع كاي عند درجة حرية 8 نجد ان

$$\lambda^* = \chi^2_{(8)}(0.05) = 15.5073$$

عليه فان المنطقة الحرجة ستكون :

$$\Sigma x_i \geq \frac{\chi^2_{(8)}(0.05)}{2\theta_0} = \frac{15.5073}{6} = 2.58455$$

او ان

$$\bar{x} \geq \frac{2.58455}{4} = 0.64614$$

عليه ترفض H_0 اذا كان متوسط العينة اكبر من او يساوي 0.64614 . وان قوة الاختبار هي :

$$1 - \beta = 1 - P_r\left(\chi^2_{(8)} \leq \left(\frac{\theta_1}{\theta_0}\right)\chi^2_{(8)}(0.05)\right)$$

$$= 1 - P_r\left(\chi^2_{(8)} \leq \frac{2}{3}(15.5073)\right)$$

$$= 1 - P_r(\chi^2_{(8)} \leq 10.3382)$$

من جداول توزيع مربع كاي نجد ان

$$P_r(\chi^2_{(8)} \leq 10.3382) \simeq 0.75$$

$$1 - \beta \simeq 0.25$$

عليه فان

مثال (٩) : افرض ان $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة عشوائية مسحوبة من $N(\theta, 1)$. جد المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار الفرضية $H_0 : \theta = \theta_0$ ضد $H_1 : \theta = \theta_1 > \theta_0$ ثم جد قوة الاختبار .

الحل : واضح ان

$$f(x,\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2}$$

عليه فان :

$$L_0 = (\sqrt{2\pi})^{-n} . e^{-\frac{1}{2}\Sigma (x_i-\theta_0)^2}$$

$$L_1 = (\sqrt{2\pi})^{-n} . e^{-\frac{1}{2}\Sigma (x_i-\theta_1)^2}$$

وحسب مبرهنة نيمان ـ پيرسون فان افضل منطقة حرجة لاختبار H_0 هي : $\frac{L_1}{L_0} \geq K$ اي ان :

$$\frac{L_1}{L_0} = e^{\frac{1}{2}\Sigma (x_i-\theta_0)^2 - \frac{1}{2}\Sigma (x_i-\theta_1)^2} \geq K$$

او ان

$$\Sigma (x_i - \theta_0)^2 - \Sigma (x_i - \theta_1)^2 \geq 2 \ln K$$

$$\Longrightarrow 2(\theta_1 - \theta_0)\Sigma x_i - n(\theta_1^2 - \theta_0^2) \geq 2 \ln K$$

$$\Longrightarrow \Sigma x_i \geq \frac{2 \ln K + n(\theta_1^2 - \theta_0^2)}{2(\theta_1 - \theta_0)}, \theta_1 - \theta_0 > 0$$

$$\Longrightarrow \bar{x} \geq \frac{2 \ln K + n(\theta_1^2 - \theta_0^2)}{2n(\theta_1 - \theta_0)} = \lambda$$

وهذا يعني ان المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار H_0 هي $\omega_0 : \{X : x \geq \lambda\}$. اي ترفض H_0 اذا كان متوسط هذه العينة اكبر من او يساوي العدد λ

الهدف الان ايجاد قيمة λ التي تحقق الصيغة :

$$\alpha = P_r (X \in \omega_0 \mid H_0)$$

اي ان

$$\alpha = P_r (\bar{x} \geq \lambda \mid H_0 : \theta = \theta_0)$$

$$\therefore \alpha = P_r \left(\frac{\bar{x} - \theta}{\sigma / \sqrt{n}} \geq \frac{\lambda - \theta}{\sigma / \sqrt{n}} \mid H_0 : \theta = \theta_0 \right) ; \sigma = 1 ,$$

$$\bar{x} \sim N \left(\theta , \frac{1}{n} \right)$$

$$= P_r (Z \geq \sqrt{n} (\lambda - \theta_0)) \quad , Z \sim N (0,1)$$

$$= 1 - P_r (Z \leq \sqrt{n} (\lambda - \theta_0))$$

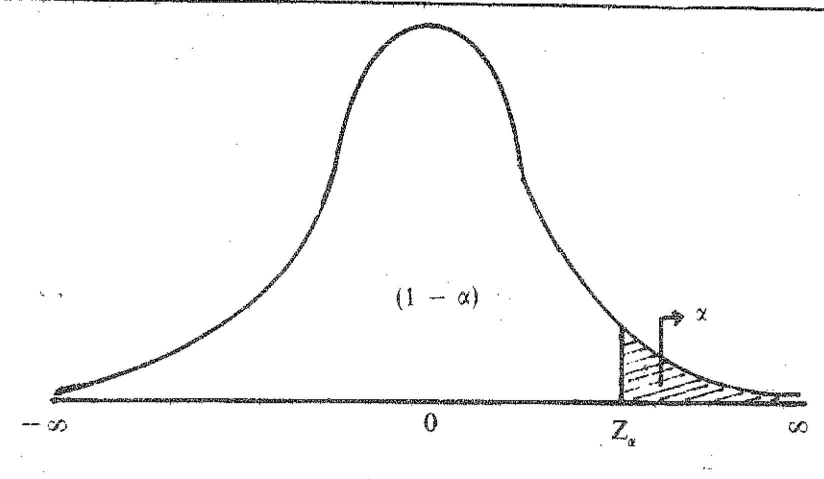
$$\therefore P_r (Z \leq \sqrt{n} (\lambda - \theta_0)) = 1 - \alpha \rightarrow \Phi (\sqrt{n} (\lambda - \theta_0))$$

$$= 1 - \alpha$$

ويتضح مما تقدم ان $\Phi (Z_0)$ ، $Z_0 = \sqrt{n} (\lambda - \theta_0) ,$ تمثل قيمة التراكم الاحتمالي في $N (0,1)$ لغاية Z_0 . وهذا يعني ان Z_0 تمثل قيمة من قيم Z التي تعطي احتمالاً متراكماً قدره $(1 - \alpha)$ والمعرفة في جداول $N (0,1)$ ولتكن Z_α . وهذا يعني ان :

$$Z_\alpha = \sqrt{n} (\lambda - \theta_0) \rightarrow \lambda = \theta_0 + \frac{Z_\alpha}{\sqrt{n}}$$

وكما هو موضح بالشكل (١٢ ـ ٤) .



الشكل (١٢ ـ ٤)

عليه فان المنطقة الحرجة الاكثر قوة لاختبار H_0 هي :

$$\omega_0 = \left\{ \bar{x} \geq \theta_0 + \frac{Z_\alpha}{\sqrt{n}} \right\}$$

وبهدف حساب قوة الاختبار فان :

$$\beta = P_r(X \in \bar{\omega}_0 \mid H_1 : \theta = \theta_1)$$

$$= P_r(\bar{x} \leq \lambda \mid H_1 : \theta = \theta_1)$$

$$= P_r(Z \leq \sqrt{n}\,(\lambda - \theta_1)),\ Z \sim N(0,1)$$

وبالتعويض عن λ بـ $\theta_0 + \frac{Z_\alpha}{\sqrt{n}}$ نحصل على :

$$\beta = P_r(Z \leq Z_\alpha - \sqrt{n}\,(\theta_1 - \theta_0))$$

$$= \Phi(Z^*),\ Z^* = Z_\alpha - \sqrt{n}\,(\theta_1 - \theta_0)$$

عليه فان قوة الاختبار هي :

$$1 - \beta = 1 - \Phi(Z^*).$$

مثال (٧) : لمعطيات المثال (٦) وبفرض ان $\theta_1 = 5, \theta_0 = 4, n = 9$ ، $\alpha = 0.05$. جد المنطقة الحرجة الاكثر قوة لاختبار H_0 ثم احسب قوة الاختبار .

الحل : من جداول $N(0,1)$ نلاحظ ان $Z_{0.05} = 1.645$. عليه فان

$$\lambda = \theta_0 + \frac{Z_{0.05}}{\sqrt{n}} = 4 + \frac{1.645}{3} = 4.5483$$

وهذا يعني ان المنطقة الحرجة هي $\omega_0 = \{\bar{x} \geq 4.5483\}$. وان قوة الاختبار هي

$$1 - \beta = 1 - \Phi(Z^*), Z^* = Z_{0.05} - \sqrt{n}\,(\theta_1 - \theta_0)$$

$$= 1.645 - 3(5 - 4) = -1.355$$

$$\therefore 1 - \beta = 1 - \Phi(-1.355)$$

من جداول $N(0,1)$ نلاحظ ان $\Phi(-1.355) \simeq 0.09$ عليه فان ،

$$1 - \beta = 0.91$$

تمارين الفصل الثاني عشر

١٢ ـ ١ : افرض ان X متغير عشوائي يتوزع كتوزيع اسي بالمعلمة θ وافرض ان المنطقة الحرجة لاختبار $H_0 : \theta = 3$ ضد $H_1 : \theta = 4$ على اساس قياسة واحدة مسحوبة من هذا التوزيع هي $x \leq 4$ جد مستوى المعنوية α وقوة الاختبار.

١٢ ـ ٢ : افرض ان X متغير عشوائي يتوزع كتوزيع ثنائي الحدين بالمعلمتين $P , n = 5$ وافرض ان المنطقة الحرجة لاختبار الفرضية $H_0 : P = 0{\cdot}50$ ضد $H_1 : P = 0{\cdot}75$ هي $x > 3$ جد حجم الاختبار وقوة هذا الاختبار.

١٢ ـ ٣ : لتكن $x_1 , x_2 , \ldots , x_n$ تمثل قياسات عينة عشوائية مسحوبة من توزيع بيتا $B(\theta , 1)$ جد المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار الفرضية $H_0 : \theta = \theta_0$ ضد $H_1 : \theta = \theta_1 < \theta_0$.

١٢ ـ ٤ : افرض ان $x_1 , x_2 , \ldots , x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع اسي بالمعلمة θ . جد المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار $H_0 : \theta = \theta_0$ ضد الفرضية $H_1 : \theta = \theta_1 > \theta_0$ ثم جد قوة الاختبار.

١٢ ـ ٥ : لمعطيات السؤال (١٢ ـ ٤) وبفرض ان $\theta_1 = 4 , \theta_0 = 3 , n = 9$ جد المنطقة الحرجة الاكثر قوة ثم احسب قوة الاختبار $\alpha = 0{\cdot}01$

١٢ ـ ٦ : افرض ان $x_1 , x_2 , \ldots , x_n$. تمثل قياسات عينة عشوائية مسحوبة من $N(0 , \sigma^2)$ جد المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار الفرضية $H_0 : \sigma^2 = \sigma_0^2$ ضد $H_1 : \sigma^2 = \sigma_1^2 > \sigma_0^2$ ثم جد قوة الاختبار.

١٢ ـ ٧ : لمعطيات السؤال (١٢ ـ ٦) وبفرض ان $\sigma_1^2 = 6 , \sigma_0^2 = 4 , n = 9$ $\alpha = 0{\cdot}05$ جد المنطقة الحرجة الاكثر قوة ثم احسب قوة الاختبار.

١٢ ـ ٨ : افرض ان $x_1 , x_2 , \ldots , x_n$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع ذي توزيع پواسون بالمعلمة λ . جد المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار $H_0 : \lambda = \lambda_0$ ضد $H_1 : \lambda = \lambda_1 > \lambda_0$ ثم جد قوة الاختبار . اذا كانت $n = 10 , \lambda_0 = 0{\cdot}1 , \lambda_1 = 0{\cdot}5 , \alpha = 0{\cdot}05$ جد تلك المنطقة الحرجة مع حساب قوة الاختبار .

١٢ ـ ٩ : اذا كانت $x_1 , x_2 , \ldots , x_n$ تمثل قياسات عينة عشوائية مسحوبة من توزيع هندسي بالمعلمة P . جد المنطقة الحرجة الاكثر قوة ذات حجم α لاختبار $H_0 : P = P_0$ ضد $H_1 : P = P_1 < P_0$ ثم جد قوة الاختبار .

١٢ ـ ١٠ : لمعطيات السؤال (١٢ ـ ٩) وبفرض ان $P_0 = 0{\cdot}5 , n = 6$, $\alpha = 0{\cdot}05 , P_1 = 0{\cdot}2$, جد المنطقة الحرجة الاكثر قوة ثم احسب قوة الاختبار .

الملاحق

ملحق (أ) المصادر

ملحق (ب) الجداول الاحصائية

ملحق (جـ) مصطلحات رياضية واحصائية

الملحق ـ أ ـ
المصادر

* * المصادر العربية * *

ـ ١ ـ د . احمد عبادة سرحان « مقدمة الاحصاء التحليلي » ، دار المعارف في مصر ، ١٩٦٢ .

٢ ـ مدني دسوقي مصطفى « مبادى في نظرية الاحتمالات والاحصاء الرياضي » ، دار النهضة العربية ،القاهرة ، ١٩٦٨ .

٣ ـ د . محمودحسن المشهداني ، امير خنا هرمز ، « الاحصاء » طبع مديرية مطبعة التعليم العالي ، الموصل ، ١٩٨٩ .

ـ ٤ ـ د . وليد النوري ، د . هلال البياتي ، د . صبري العاني ، « الاحصاء الرياضي » ، طبع مديرية مطبعة جامعة الموصل ، ١٩٨٢ .

** المصادر الاجنبية **

1- AMIR H. HERMIZ: "The distribution of absolute standard normal variate", Jornal of Tanmiat Al- Rafidian, 30, Mosul, Iraq. 1990.

2- EDWARD J. DUDEWICZ: "Introduction to statistics and probability", Holt, Rinehart and Winston, New York, 1976.

3- GUPTA S. C., V. K. KAPOOR: "Fundementals of mathematical statistics", Sultan chand & Sons publishers, New Delhi, 1982.

4- HOGG V. R., A. T. CRAIG, "Introduction to mathematical statistics", Macmillan publishing Co. Ink, New York, 1970.

5- HAROLD J. LARSON: "Introduction to probability theory and statistical inference", 3rd ed., John Wiley & Sons, Ink, New York, 1982.

6- IAN F. BLAKE: "An introduction to applied probability", John Wiley & Sons, Ink, New York, 1979.

7- JOHNSON and KOTZ: "Discrete distributions", John Wiley & Sons, Ink, New York, 1969.

8- JOHNSON and KOTZ: "Continuous univariate distributions-1", John Wiley & Sons, Ink, New York, 1970.

9- JOHNSON and KOTZ: "Continuous univariate distributions-2", John Wiley & Sons, Ink, New York, 1970.

10- KAPUR J. N., H. C. SAXENA: "Mathematical statistiscs", chand &Co. (pvt.) LTD, New Delhi, 1972.

11- LUKACS E, R. G. LAHA: "Application of characteristic functions", charles Griffin & company limited, London, 1964.

12- MICHAEL WOODROOFE: "Probability with application", Mc Graw-Hill book Co., New York, 1975.

13- MOOD A. M., F. A. GRAYBILL: "Introduction to the theory of statistics", Mc Graw-Hill book Co., New York, 1963.

14- MOOD A. M., F. A. GRAYBILL and DUANE C. B.: "Introduction to the theory of statistics", Mc Graw-Hill book Co., New York, 1974.

15- MORAN P. A. P.: "Calculation of the normal distribution function", Biometrica 67, pp. 675-6, 1980.

16- PAUL G. HOEL: "Introduction to mathematical statistics", John Wiley & Sons, Ink, New York, 1970.

الملحق ـ ب ـ

* * * جداول احصائية * * *

١ ـ جداول توزيع ثنائي الحدين .

٢ ـ جداول التوزيع الهندسي الزائدي .

٣ ـ جداول توزيع پواسون .

٤ ـ جداول التوزيع الطبيعي

٥ ـ جداول توزيع بيتا

٦ ـ جداول توزيع مربع كاي

٧ ـ جداول توزيع t

٨ ـ جداول توزيع F

٩ ـ جداول توزيع المدى القياسي

الجدول (١ ـ ب) : توزيع ثنائي الحدين

$$P(x) = C_x^n \, P^x (1 - P)^{n-x}$$

$$x = 0, 1, 2, \ldots, n$$

$$0 < P < 1$$

		p									
n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
1	0	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500	.5000
	1	.0500	.1000	.1600	.2000	.2500	.3000	.3500	.4000	.4500	.5000
2	0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500
	1	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	.5000
	2	.0025	.0100	.0225	.0400	.0625	.0900	.1225	.1600	.2025	.2500
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250
4	0	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915	.0625
	1	.1715	.2916	.3685	.4096	.4219	.4116	.3845	.3456	.2995	.2500
	2	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3675	.3750
	3	.0005	.0036	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500
	4	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	.0625
5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0312
	1	.2036	.3280	.3915	.4096	.3955	.3602	.3124	.2592	.2059	.1562
	2	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125
	3	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3125
	4	.0000	.0004	.0022	.0064	.0146	.0284	.0488	.0768	.1128	.1562
	5	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	.0312
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
	1	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938
	2	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	.2344
	3	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3125
	4	.0001	.0012	.0055	.0154	.0330	.0595	.0951	.1382	.1861	.2344
	5	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938
	6	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	.0156
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0079
	1	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547
	2	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641
	3	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734
	4	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734
	5	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039
	1	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0312
	2	.0515	.1488	.2376	.2936	.3115	.2965	.2587	.2090	.1569	.1094
	3	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188
	4	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734
	5	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094
	7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039

تابع الجدول (١ – ب)

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
						p					
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098
	2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172
	4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
11	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005
	1	.3293	.3835	.3248	.2362	.1549	.0932	.0518	.0266	.0125	.0054
	2	.0867	.2131	.2866	.2953	.2581	.1998	.1395	.0887	.0513	.0269
	3	.0137	.0710	.1517	.2215	.2581	.2568	.2254	.1774	.1259	.0806
	4	.0014	.0158	.0536	.1107	.1721	.2201	.2428	.2365	.2060	.1611
	5	.0001	.0025	.0132	.0388	.0803	.1321	.1830	.2207	.2360	.2256
	6	.0000	.0003	.0023	.0097	.0268	.0566	.0985	.1471	.1931	.2256
	7	.0000	.0000	.0003	.0017	.0064	.0173	.0379	.0701	.1128	.1611
	8	.0000	.0000	.0000	.0002	.0011	.0037	.0102	.0234	.0462	.0806
	9	.0000	.0000	.0000	.0000	.0001	.0005	.0018	.0052	.0126	.0269
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0007	.0021	.0054
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0005
12	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029
	2	.0988	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161
	3	.0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537
	4	.0021	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208
	5	.0002	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934
	6	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256
	7	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934
	8	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208
	9	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0537
	10	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016
	2	.1109	.2448	.2937	.2680	.2059	.1388	.0836	.0453	.0220	.0095
	3	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349
	4	.0028	.0277	.0838	.1535	.2097	.2337	.2222	.1845	.1350	.0873
	5	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571
	6	.0000	.0008	.0063	.0230	.0559	.1030	.1546	.1968	.2169	.2095
	7	.0000	.0001	.0011	.0058	.0180	.0442	.0833	.1312	.1775	.2095
	8	.0000	.0000	.0001	.0011	.0047	.0142	.0336	.0656	.1089	.1571
	9	.0000	.0000	.0000	.0001	.0009	.0034	.0101	.0243	.0495	.0873
	10	.0000	.0000	.0000	.0000	.0001	.0006	.0022	.0065	.0162	.0349
	11	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012	.0036	.0095
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

تابع الجدول (١ ـ ب)

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
						p					
14	0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001
	1	.3593	.3559	.2539	.1539	.0832	.0407	.0181	.0073	.0027	.0009
	2	.1229	.2570	.2912	.2501	.1802	.1134	.0634	.0317	.0141	.0056
	3	.0259	.1142	.2056	.2501	.2402	.1943	.1366	.0845	.0462	.0222
	4	.0037	.0348	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611
	5	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222
	6	.0000	.0013	.0093	.0322	.0734	.1262	.1759	.2066	.2088	.1833
	7	.0000	.0002	.0019	.0092	.0280	.0618	.1082	.1574	.1952	.2095
	8	.0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833
	9	.0000	.0000	.0000	.0003	.0018	.0066	.0183	.0408	.0762	.1222
	10	.0000	.0000	.0000	.0000	.0003	.0014	.0049	.0136	.0312	.0611
	11	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	.0093	.0222
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0056
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0009
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
15	0	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	.0000
	1	.3658	.3432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005
	2	.1348	.2669	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032
	3	.0307	.1285	.2184	.2501	.2252	.1700	.1110	.0634	.0318	.0139
	4	.0049	.0428	.1156	.1876	.2252	.2186	.1792	.1268	.0780	.0417
	5	.0006	.0105	.0449	.1032	.1651	.2061	.2123	.1859	.1404	.0916
	6	.0000	.0019	.0132	.0430	.0917	.1472	.1906	.2066	.1914	.1527
	7	.0000	.0003	.0030	.0138	.0393	.0811	.1319	.1771	.2013	.1964
	8	.0000	.0000	.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964
	9	.0000	.0000	.0001	.0007	.0034	.0116	.0298	.0612	.1048	.1527
	10	.0000	.0000	.0000	.0001	.0007	.0030	.0090	.0245	.0515	.0916
	11	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0074	.0191	.0417
	12	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	.0139
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0032
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000
	1	.3706	.3294	.2097	.1126	.0535	.0228	.0087	.0030	.0009	.0002
	2	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018
	3	.0359	.1423	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085
	4	.0061	.0514	.1311	.2001	.2252	.2040	.1553	.1014	.0572	.0278
	5	.0008	.0137	.0555	.1201	.1802	.2099	.2008	.1623	.1123	.0667
	6	.0001	.0028	.0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222
	7	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746
	8	.0000	.0001	.0009	.0055	.0197	.0487	.0923	.1417	.1812	.1964
	9	.0000	.0000	.0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746
	10	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222
	11	.0000	.0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667
	12	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0029	.0085
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
17	0	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000	.0000
	1	.3741	.3150	.1893	.0957	.0426	.0169	.0060	.0019	.0005	.0001
	2	.1575	.2800	.2673	.1914	.1136	.0581	.0260	.0102	.0035	.0010
	3	.0415	.1556	.2359	.2393	.1893	.1245	.0701	.0341	.0144	.0052
	4	.0076	.0605	.1457	.2093	.2209	.1868	.1320	.0796	.0411	.0182
	5	.0010	.0175	.0668	.1361	.1914	.2081	.1849	.1379	.0875	.0472
	6	.0001	.0039	.0236	.0680	.1276	.1784	.1991	.1839	.1432	.0944
	7	.0000	.0007	.0065	.0267	.0668	.1201	.1685	.1927	.1841	.1484
	8	.0000	.0001	.0014	.0084	.0279	.0644	.1134	.1606	.1883	.1855
	9	.0000	.0000	.0003	.0021	.0093	.0276	.0611	.1070	.1540	.1855
	10	.0000	.0000	.0000	.0004	.0025	.0095	.0263	.0571	.1008	.1484
	11	.0000	.0000	.0000	.0001	.0005	.0026	.0090	.0242	.0525	.0944
	12	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0081	.0215	.0472
	13	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0021	.0068	.0182
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052

(تابع الجدول (١ ـ ب)

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
17	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
18	0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000
	1	.3763	.3002	.1704	.0811	.0338	.0126	.0042	.0012	.0003	.0001
	2	.1683	.2835	.2556	.1723	.0958	.0458	.0190	.0069	.0022	.0006
	3	.0473	.1680	.2406	.2297	.1704	.1046	.0547	.0246	.0095	.0031
	4	.0093	.0700	.1592	.2153	.2130	.1681	.1104	.0614	.0291	.0117
	5	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327
	6	.0002	.0052	.0301	.0816	.1436	.1873	.1941	.1655	.1181	.0708
	7	.0000	.0010	.0091	.0350	.0820	.1376	.1792	.1892	.1657	.1214
	8	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1864	.1669
	9	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1694	.1855
	10	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669
	11	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214
	12	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708
	13	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0045	.0134	.0327
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
19	0	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000
	1	.3774	.2852	.1529	.0685	.0268	.0093	.0029	.0008	.0002	.0000
	2	.1787	.2852	.2428	.1540	.0803	.0358	.0138	.0046	.0013	.0003
	3	.0533	.1796	.2428	.2182	.1517	.0869	.0422	.0175	.0062	.0018
	4	.0112	.0798	.1714	.2182	.2023	.1491	.0909	.0467	.0203	.0074
	5	.0018	.0266	.0907	.1636	.2023	.1916	.1468	.0933	.0497	.0222
	6	.0002	.0069	.0374	.0955	.1574	.1916	.1844	.1451	.0949	.0518
	7	.0000	.0014	.0122	.0443	.0974	.1525	.1844	.1797	.1443	.0961
	8	.0000	.0002	.0032	.0166	.0487	.0981	.1489	.1797	.1771	.1442
	9	.0000	.0000	.0007	.0051	.0198	.0514	.0980	.1464	.1771	.1762
	10	.0000	.0000	.0001	.0013	.0066	.0220	.0528	.0976	.1449	.1762
	11	.0000	.0000	.0000	.0003	.0018	.0077	.0233	.0532	.0970	.1442
	12	.0000	.0000	.0000	.0000	.0004	.0022	.0083	.0237	.0529	.0961
	13	.0000	.0000	.0000	.0000	.0001	.0005	.0024	.0085	.0233	.0518
	14	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0082	.0222
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0022	.0074
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
20	0	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000
	1	.3774	.2702	.1368	.0576	.0211	.0068	.0020	.0005	.0001	.0000
	2	.1887	.2852	.2293	.1369	.0669	.0278	.0100	.0031	.0008	.0002
	3	.0596	.1901	.2428	.2054	.1339	.0716	.0323	.0123	.0040	.0011
	4	.0133	.0898	.1821	.2182	.1897	.1304	.0738	.0350	.0139	.0046
	5	.0022	.0319	.1028	.1746	.2023	.1789	.1272	.0746	.0365	.0148
	6	.0003	.0089	.0454	.1091	.1686	.1916	.1712	.1244	.0746	.0370
	7	.0000	.0020	.0160	.0545	.1124	.1643	.1844	.1659	.1221	.0739
	8	.0000	.0004	.0046	.0222	.0609	.1144	.1614	.1797	.1623	.1201
	9	.0000	.0001	.0011	.0074	.0271	.0654	.1158	.1597	.1771	.1602
	10	.0000	.0000	.0002	.0020	.0099	.0308	.0686	.1171	.1593	.1762
	11	.0000	.0000	.0000	.0005	.0030	.0120	.0336	.0710	.1185	.1602
	12	.0000	.0000	.0000	.0001	.0008	.0039	.0136	.0355	.0727	.1201
	13	.0000	.0000	.0000	.0000	.0002	.0010	.0045	.0146	.0366	.0739
	14	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0049	.0150	.0370
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0049	.0148
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0046
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

الجدول (٢ ـ ب) : التوزيع الهندسي الزائدي

$$P(x) = \frac{C_x^M \cdot C_{n-x}^{N-M}}{C_n^N}$$

$$\max.(0, n - N + M) \le x \le \min.(n, M) \text{ ، } a = M$$

N	n	a	x	$p_X(x)$
2	1	1	0	.500000
2	1	1	1	.500000
3	1	1	0	.666667
3	1	1	1	.333333
3	2	1	0	.333333
3	2	1	1	.666667
3	2	2	1	.666667
3	2	2	2	.333333
4	1	1	0	.750000
4	1	1	1	.250000
4	2	1	0	.500000
4	2	1	1	.500000
4	2	2	0	.166667
4	2	2	1	.666667
4	2	2	2	.166667
4	3	1	0	.250000
4	3	1	1	.750000
4	3	2	1	.500000
4	3	2	2	.500000
4	3	3	2	.750000
4	3	3	3	.250000
5	1	1	0	.800000
5	1	1	1	.200000
5	2	1	0	.600000
5	2	1	1	.400000
5	2	2	0	.300000
5	2	2	1	.600000
5	2	2	2	.100000
5	3	1	0	.400000
5	3	1	1	.600000
5	3	2	0	.100000
5	3	2	1	.600000
5	3	2	2	.300000
5	3	3	1	.300000
5	3	3	2	.600000
5	3	3	3	.100000
5	4	1	0	.200000
5	4	1	1	.800000
5	4	2	1	.400000
5	4	2	2	.600000

N	n	a	x	$p_X(x)$
5	4	3	2	.600000
5	4	3	3	.400000
5	4	4	3	.800000
5	4	4	4	.200000
6	1	1	0	.833333
6	1	1	1	.166667
6	2	1	0	.666667
6	2	1	1	.333333
6	2	2	0	.400000
6	2	2	1	.533333
6	2	2	2	.066667
6	3	1	0	.500000
6	3	1	1	.500000
6	3	2	0	.200000
6	3	2	1	.600000
6	3	2	2	.200000
6	3	3	0	.050000
6	3	3	1	.450000
6	3	3	2	.450000
6	3	3	3	.050000
6	4	1	0	.333333
6	4	1	1	.666667
6	4	2	0	.066667
6	4	2	1	.533333
6	4	2	2	.400000
6	4	3	1	.200000
6	4	3	2	.600000
6	4	3	3	.200000
6	4	4	2	.400000
6	4	4	3	.533333
6	4	4	4	.066667
6	5	1	0	.166667
6	5	1	1	.833333
6	5	2	1	.333333
6	5	2	2	.666667
6	5	3	2	.500000
6	5	3	3	.500000
6	5	4	3	.666667
6	5	4	4	.333333
6	5	5	4	.833333

N	n	a	x	$p_X(x)$
6	5	5	5	.166667
7	1	1	0	.857143
7	1	1	1	.142857
7	2	1	0	.714286
7	2	1	1	.285714
7	2	2	0	.476190
7	2	2	1	.476190
7	2	2	2	.047619
7	3	1	0	.571429
7	3	1	1	.428571
7	3	2	0	.285714
7	3	2	1	.571429
7	3	2	2	.142857
7	3	3	0	.114286
7	3	3	1	.514286
7	3	3	2	.342857
7	3	3	3	.028571
7	4	1	0	.428571
7	4	1	1	.571429
7	4	2	0	.142857
7	4	2	1	.571429
7	4	2	2	.285714
7	4	3	0	.028571
7	4	3	1	.342857
7	4	3	2	.514286
7	4	3	3	.114286
7	4	4	1	.114286
7	4	4	2	.514286
7	4	4	3	.342857
7	4	4	4	.028571
7	5	1	0	.285714
7	5	1	1	.714286
7	5	2	0	.047619
7	5	2	1	.476190
7	5	2	2	.476190
7	5	3	1	.142857
7	5	3	2	.571429
7	5	3	3	.285714
7	5	4	2	.285714
7	5	4	3	.571429

N	n	a	x	$p_X(x)$
7	5	4	4	.142857
7	5	5	3	.476190
7	5	5	4	.476190
7	5	5	5	.047619
7	6	1	0	.142857
7	6	1	1	.857143
7	6	2	1	.285714
7	6	2	2	.714286
7	6	3	2	.428571
7	6	3	3	.571429
7	6	4	3	.571429
7	6	4	4	.428571
7	6	5	4	.714286
7	6	5	5	.285714
7	6	6	5	.857143
7	6	6	6	.142857
8	1	1	0	.875000
8	1	1	1	.125000
8	2	1	0	.750000
8	2	1	1	.250000
8	2	2	0	.535714
8	2	2	1	.428571
8	2	2	2	.035714
8	3	1	0	.625000
8	3	1	1	.375000
8	3	2	0	.357143
8	3	2	1	.535714
8	3	2	2	.107143
8	3	3	0	.178571
8	3	3	1	.535714
8	3	3	2	.267857
8	3	3	3	.017857
8	4	1	0	.500000
8	4	1	1	.500000
8	4	2	0	.214286
8	4	2	1	.571429
8	4	2	2	.214286
8	4	3	0	.071429
8	4	3	1	.428571
8	4	3	2	.428571

تابع الجدول (٢ – ب)

N	n	a	x	$p_X(x)$	N	n	a	x	$p_X(x)$	N	n	a	x	$p_X(x)$	N	n	a	x	$p_X(x)$
8	4	3	3	.071429	8	7	1	0	.125000	9	4	4	1	.317460	9	6	6	3	.238095
8	4	4	0	.014286	8	7	1	1	.875000	9	4	4	2	.476190	9	6	6	4	.535714
8	4	4	1	.228571	8	7	2	1	.250000	9	4	4	3	.158730	9	6	6	5	.214286
8	4	4	2	.514286	8	7	2	2	.750000	9	4	4	4	.007936	9	6	6	6	.011905
8	4	4	3	.228571	8	7	3	2	.375000	9	5	1	0	.444444	9	7	1	0	.222222
8	4	4	4	.014286	8	7	3	3	.625000	9	5	1	1	.555556	9	7	1	1	.777778
8	5	1	0	.375000	8	7	4	3	.500000	9	5	2	0	.166667	9	7	2	0	.027778
8	5	1	1	.625000	8	7	4	4	.500000	9	5	2	1	.555556	9	7	2	1	.388889
8	5	2	0	.107143	8	7	5	4	.625000	9	5	2	2	.277778	9	7	2	2	.583333
8	5	2	1	.535714	8	7	5	5	.375000	9	5	3	0	.047619	9	7	3	1	.083333
8	5	2	2	.357143	8	7	6	5	.750000	9	5	3	1	.357143	9	7	3	2	.500000
8	5	3	0	.017857	8	7	6	6	.250000	9	5	3	2	.476190	9	7	3	3	.416667
8	5	3	1	.267857	8	7	7	6	.875000	9	5	3	3	.119048	9	7	4	2	.166667
8	5	3	2	.535714	8	7	7	7	.125000	9	5	4	0	.007936	9	7	4	3	.555556
8	5	3	3	.178571	9	1	1	0	.888889	9	5	4	1	.158730	9	7	4	4	.277778
8	5	4	1	.071429	9	1	1	1	.111111	9	5	4	2	.476190	9	7	5	3	.277778
8	5	4	2	.428571	9	2	1	0	.777778	9	5	4	3	.317460	9	7	5	4	.555556
8	5	4	3	.428571	9	2	1	1	.222222	9	5	4	4	.039683	9	7	5	5	.166667
8	5	4	4	.071429	9	2	2	0	.583333	9	5	5	1	.039683	9	7	6	4	.416667
8	5	5	2	.178571	9	2	2	1	.388889	9	5	5	2	.317460	9	7	6	5	.500000
8	5	5	3	.535714	9	2	2	2	.027778	9	5	5	3	.476190	9	7	6	6	.083333
8	5	5	4	.267857	9	3	1	0	.666667	9	5	5	4	.158730	9	7	7	5	.583333
8	5	5	5	.017857	9	3	1	1	.333333	9	5	5	5	.007936	9	7	7	6	.388889
8	6	1	0	.250000	9	3	2	0	.416667	9	6	1	0	.333333	9	7	7	7	.027778
8	6	1	1	.750000	9	3	2	1	.500000	9	6	1	1	.666667	9	8	1	0	.111111
8	6	2	0	.015714	9	3	2	2	.083333	9	6	2	0	.083333	9	8	1	1	.888889
8	6	2	1	.428571	9	3	3	0	.238095	9	6	2	1	.500000	9	8	2	1	.222222
8	6	2	2	.535714	9	3	3	1	.535714	9	6	2	2	.416667	9	8	2	2	.777778
8	6	3	1	.107143	9	3	3	2	.214286	9	6	3	0	.011905	9	8	3	2	.333333
8	6	3	2	.535714	9	3	3	3	.011905	9	6	3	1	.214286	9	8	3	3	.666667
8	6	3	3	.357143	9	4	1	0	.555556	9	6	3	2	.535714	9	8	4	3	.444444
8	6	4	2	.214286	9	4	1	1	.444444	9	6	3	3	.238095	9	8	4	4	.555556
8	6	4	3	.571429	9	4	2	0	.277778	9	6	4	1	.047619	9	8	5	4	.555556
8	6	4	4	.214286	9	4	2	1	.555556	9	6	4	2	.357143	9	8	5	5	.444444
8	6	5	3	.357143	9	4	2	2	.166667	9	6	4	3	.476190	9	8	6	5	.666667
8	6	5	4	.535714	9	4	3	0	.119048	9	6	4	4	.119048	9	8	6	6	.333333
8	6	5	5	.107143	9	4	3	1	.476190	9	6	5	2	.119048	9	8	7	6	.777778
8	6	6	4	.535714	9	4	3	2	.357143	9	6	5	3	.476190	9	8	7	7	.222222
8	6	6	5	.428571	9	4	3	3	.047619	9	6	5	4	.357143	9	8	8	7	.888889
8	6	6	6	.035714	9	4	4	0	.039683	9	6	5	5	.047619	9	8	8	8	.111111

تابع الجدول (٢ ـ ب)

N	n	a	x	$p_X(x)$
10	1	1	0	.900000
10	1	1	1	.100000
10	2	1	0	.800000
10	2	1	1	.200000
10	2	2	0	.622222
10	2	2	1	.355556
10	2	2	2	.022222
10	3	1	0	.700000
10	3	1	1	.300000
10	3	2	0	.466667
10	3	2	1	.466667
10	3	2	2	.066667
10	3	3	0	.291667
10	3	3	1	.525000
10	3	3	2	.175000
10	3	3	3	.008333
10	4	1	0	.600000
10	4	1	1	.400000
10	4	2	0	.333333
10	4	2	1	.533333
10	4	2	2	.133333
10	4	3	0	.166667
10	4	3	1	.500000
10	4	3	2	.300000
10	4	3	3	.033333
10	4	4	0	.071429
10	4	4	1	.380952
10	4	4	2	.428571
10	4	4	3	.114286
10	4	4	4	.004762
10	5	1	0	.500000
10	5	1	1	.500000
10	5	2	0	.222222
10	5	2	1	.555556
10	5	2	2	.222222
10	5	3	0	.083333
10	5	3	1	.416667
10	5	3	2	.416667
10	5	3	3	.083333
10	5	4	0	.023810
10	5	4	1	.238095
10	5	4	2	.476190
10	5	4	3	.238095
10	5	4	4	.023810
10	5	5	0	.003968
10	5	5	1	.099206
10	5	5	2	.396825
10	5	5	3	.396825
10	5	5	4	.099206
10	5	5	5	.003968
10	6	1	0	.400000
10	6	1	1	.600000
10	6	2	0	.133333
10	6	2	1	.533333
10	6	2	2	.333333
10	6	3	0	.033333
10	6	3	1	.300000
10	6	3	2	.500000
10	6	3	3	.166667
10	6	4	0	.004762
10	6	4	1	.114286
10	6	4	2	.428571
10	6	4	3	.380952
10	6	4	4	.071429
10	6	5	1	.023810
10	6	5	2	.238095
10	6	5	3	.476190
10	6	5	4	.238095
10	6	5	5	.023810
10	6	6	2	.071429
10	6	6	3	.380952
10	6	6	4	.428571
10	6	6	5	.114286
10	6	6	6	.004762
10	7	1	0	.300000
10	7	1	1	.700000
10	7	2	0	.066667
10	7	2	1	.466667
10	7	2	2	.466667
10	7	3	0	.008333
10	7	3	1	.175000
10	7	3	2	.525000
10	7	3	3	.291667
10	7	4	1	.033333
10	7	4	2	.300000
10	7	4	3	.500000
10	7	4	4	.166667
10	7	5	2	.083333
10	7	5	3	.416667
10	7	5	4	.416667
10	7	5	5	.083333
10	7	6	3	.166667
10	7	6	4	.500000
10	7	6	5	.300000
10	7	6	6	.033333
10	7	7	4	.291667
10	7	7	5	.525000
10	7	7	6	.175000
10	7	7	7	.008333
10	8	1	0	.200000
10	8	1	1	.800000
10	8	2	0	.022222
10	8	2	1	.355556
10	8	2	2	.622222
10	8	3	1	.066667
10	8	3	2	.466667
10	8	3	3	.466667
10	8	4	2	.133333
10	8	4	3	.533333
10	8	4	4	.333333
10	8	5	3	.222222
10	8	5	4	.555556
10	8	5	5	.222222
10	8	6	4	.333333
10	8	6	5	.533333
10	8	6	6	.133333
10	8	7	5	.466667
10	8	7	6	.466667
10	8	7	7	.066667
10	8	8	6	.622222
10	8	8	7	.355556
10	8	8	8	.022222
10	9	1	0	.100000
10	9	1	1	.900000
10	9	2	1	.200000
10	9	2	2	.800000
10	9	3	2	.300000
10	9	3	3	.700000
10	9	4	3	.400000
10	9	4	4	.600000
10	9	5	4	.500000
10	9	5	5	.500000
10	9	6	5	.600000
10	9	6	6	.400000
10	9	7	6	.700000
10	9	7	7	.300000
10	9	8	7	.800000
10	9	8	8	.200000
10	9	9	8	.900000
10	9	9	9	.100000

الجدول (٢ ـ ب) : توزيع بواسون

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \;; x = 0, 1, 2, \ldots, \mu = \lambda$$

x	λ .1	.2	.3	4	.5	.6	.7	.8	.9	1.0
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.0905	.1637	.2222	.2681	.3033	.3293	.3476	.3595	.3659	.3679
2	.0045	.0164	.0333	.0536	.0758	.0988	.1217	.1438	.1647	.1839
3	.0002	.0011	.0033	.0072	.0126	.0198	.0284	.0383	.0494	.0613
4	.0000	.0001	.0002	.0007	.0016	.0030	.0050	.0077	.0111	.0153
5	.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0012	.0020	.0031
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x	λ 1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	.3329	.3012	.2725	.2466	.2231	.2019	.1827	.1653	.1496	.1353
1	.3662	.3614	.3543	.3452	.3347	.3230	.3106	.2975	.2842	.2707
2	.2014	.2169	.2303	.2417	.2510	.2584	.2640	.2678	.2700	.2707
3	.0738	.0867	.0998	.1128	.1255	.1378	.1496	.1607	.1710	.1804
4	.0203	.0260	.0324	.0395	.0471	.0551	.0636	.0723	.0812	.0902
5	.0045	.0062	.0084	.0111	.0141	.0176	.0216	.0260	.0309	.0361
6	.0008	.0012	.0018	.0026	.0035	.0047	.0061	.0078	.0098	.0120
7	.0001	.0002	.0003	.0005	.0008	.0011	.0015	.0020	.0027	.0034
8	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0005	.0006	.0009
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

x	λ 2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	.0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081
9	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018	.0022	.0027
10	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x	λ 3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
0	.0450	.0408	.0369	.0334	.0302	.0273	.0247	.0224	.0202	.0183
1	.1397	.1304	.1217	.1135	.1057	.0984	.0915	.0850	.0789	.0733
2	.2165	.2087	.2008	.1929	.1850	.1771	.1692	.1615	.1539	.1465
3	.2237	.2226	.2209	.2186	.2158	.2125	.2087	.2046	.2001	.1954
4	.1734	.1781	.1823	.1858	.1888	.1912	.1931	.1944	.1951	.1954
5	.1075	.1140	.1203	.1264	.1322	.1377	.1429	.1477	.1522	.1563
6	.0555	.0608	.0662	.0716	.0771	.0826	.0881	.0936	.0989	.1042
7	.0246	.0278	.0312	.0348	.0385	.0425	.0466	.0508	.0551	.0595
8	.0095	.0111	.0129	.0148	.0169	.0191	.0215	.0241	.0269	.0298
9	.0033	.0040	.0047	.0056	.0066	.0076	.0089	.0102	.0116	.0132
10	.0010	.0013	.0016	.0019	.0023	.0028	.0033	.0039	.0045	.0053
11	.0003	.0004	.0005	.0006	.0007	.0009	.0011	.0013	.0016	.0019
12	.0001	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006
13	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

تابع الجدول (٢ ـ ب)

x	λ 4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
0	.0166	.0150	.0136	.0123	.0111	.0101	.0091	.0082	.0074	.0067
1	.0679	.0630	.0583	.0540	.0500	.0462	.0427	.0395	.0365	.0337
2	.1393	.1323	.1254	.1188	.1125	.1063	.1005	.0948	.0894	.0842
3	.1904	.1852	.1798	.1743	.1687	.1631	.1574	.1517	.1460	.1404
4	.1951	.1944	.1933	.1917	.1898	.1875	.1849	.1820	.1789	.1755
5	.1600	.1633	.1662	.1687	.1708	.1725	.1738	.1747	.1753	.1755
6	.1093	.1143	.1191	.1237	.1281	.1323	.1362	.1398	.1432	.1462
7	.0640	.0686	.0732	.0778	.0824	.0869	.0914	.0959	.1002	.1044
8	.0328	.0360	.0393	.0428	.0463	.0500	.0537	.0575	.0614	.0653
9	.0150	.0168	.0188	.0209	.0232	.0255	.0280	.0307	.0334	.0363
10	.0061	.0071	.0081	.0092	.0104	.0118	.0132	.0147	.0164	.0181
11	.0023	.0027	.0032	.0037	.0043	.0049	.0056	.0064	.0073	.0082
12	.0008	.0009	.0011	.0014	.0016	.0019	.0022	.0026	.0030	.0034
13	.0002	.0003	.0004	.0005	.0006	.0007	.0008	.0009	.0011	.0013
14	.0001	.0001	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005
15	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0002

x	λ 5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
0	.0061	.0055	.0050	.0045	.0041	.0037	.0033	.0030	.0027	.0025
1	.0311	.0287	.0265	.0244	.0225	.0207	.0191	.0176	.0162	.0149
2	.0793	.0746	.0701	.0659	.0618	.0580	.0544	.0509	.0477	.0446
3	.1348	.1293	.1239	.1185	.1133	.1082	.1033	.0985	.0938	.0892
4	.1719	.1681	.1641	.1600	.1558	.1515	.1472	.1428	.1383	.1339
5	.1753	.1748	.1740	.1728	.1714	.1697	.1678	.1656	.1632	.1606
6	.1490	.1515	.1537	.1555	.1571	.1584	.1594	.1601	.1605	.1606
7	.1086	.1125	.1163	.1200	.1234	.1267	.1298	.1326	.1353	.1377
8	.0692	.0731	.0771	.0810	.0849	.0887	.0925	.0962	.0998	.1033
9	.0392	.0423	.0454	.0486	.0519	.0552	.0586	.0620	.0654	.0688
10	.0200	.0220	.0241	.0262	.0285	.0309	.0334	.0359	.0386	.0413
11	.0093	.0104	.0116	.0129	.0143	.0157	.0173	.0190	.0207	.0225
12	.0039	.0045	.0051	.0058	.0065	.0073	.0082	.0092	.0102	.0113
13	.0015	.0018	.0021	.0024	.0028	.0032	.0036	.0041	.0046	.0052
14	.0006	.0007	.0008	.0009	.0011	.0013	.0015	.0017	.0019	.0022
15	.0002	.0002	.0003	.0003	.0004	.0005	.0006	.0007	.0008	.0009
16	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003
17	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001

x	λ 6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
0	.0022	.0020	.0018	.0017	.0015	.0014	.0012	.0011	.0010	.0009
1	.0137	.0126	.0116	.0106	.0098	.0090	.0082	.0076	.0070	.0064
2	.0417	.0390	.0364	.0340	.0318	.0296	.0276	.0258	.0240	.0223
3	.0848	.0806	.0765	.0726	.0688	.0652	.0617	.0584	.0552	.0521
4	.1294	.1249	.1205	.1162	.1118	.1076	.1034	.0992	.0952	.0912
5	.1579	.1549	.1519	.1487	.1454	.1420	.1385	.1349	.1314	.1277
6	.1605	.1601	.1595	.1586	.1575	.1562	.1546	.1529	.1511	.1490
7	.1399	.1418	.1435	.1450	.1462	.1472	.1480	.1486	.1489	.1490
8	.1066	.1099	.1130	.1160	.1188	.1215	.1240	.1263	.1284	.1304
9	.0723	.0757	.0791	.0825	.0858	.0891	.0923	.0954	.0985	.1014
10	.0441	.0469	.0498	.0528	.0558	.0588	.0618	.0649	.0679	.0710
11	.0245	.0265	.0285	.0307	.0330	.0353	.0377	.0401	.0426	.0452
12	.0124	.0137	.0150	.0164	.0179	.0194	.0210	.0227	.0245	.0264
13	.0058	.0065	.0073	.0081	.0089	.0098	.0108	.0119	.0130	.0142
14	.0025	.0029	.0033	.0037	.0041	.0046	.0052	.0058	.0064	.0071
15	.0010	.0012	.0014	.0016	.0018	.0020	.0023	.0026	.0029	.0033
16	.0004	.0005	.0005	.0006	.0007	.0008	.0010	.0011	.0013	.0014
17	.0001	.0002	.0002	.0002	.0003	.0003	.0004	.0004	.0005	.0006
18	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002
19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

تابع الجدول (٢ ـ ب)

x	λ 7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	.0008	.0007	.0007	.0006	.0006	.0005	.0005	.0004	.0004	.0003
1	.0059	.0054	.0049	.0045	.0041	.0038	.0035	.0032	.0029	.0027
2	.0208	.0194	.0180	.0167	.0156	.0145	.0134	.0125	.0116	.0107
3	.0492	.0464	.0438	.0413	.0389	.0366	.0345	.0324	.0305	.0286
4	.0874	.0836	.0799	.0764	.0729	.0696	.0663	.0632	.0602	.0573
5	.1241	.1204	.1167	.1130	.1094	.1057	.1021	.0986	.0951	.0916
6	.1468	.1445	.1420	.1394	.1367	.1339	.1311	.1282	.1252	.1221
7	.1489	.1486	.1481	.1474	.1465	.1454	.1442	.1428	.1413	.1396
8	.1321	.1337	.1351	.1363	.1373	.1382	.1388	.1392	.1395	.1396
9	.1042	.1070	.1096	.1121	.1144	.1167	.1187	.1207	.1224	.1241
10	.0740	.0770	.0800	.0829	.0858	.0887	.0914	.0941	.0967	.0993
11	.0478	.0504	.0531	.0558	.0585	.0613	.0640	.0667	.0695	.0722
12	.0283	.0303	.0323	.0344	.0366	.0388	.0411	.0434	.0457	.0481
13	.0154	.0168	.0181	.0196	.0211	.0227	.0243	.0260	.0278	.0296
14	.0078	.0086	.0095	.0104	.0113	.0123	.0134	.0145	.0157	.0169
15	.0037	.0041	.0046	.0051	.0057	.0062	.0069	.0075	.0083	.0090
16	.0016	.0019	.0021	.0024	.0026	.0030	.0033	.0037	.0041	.0045
17	.0007	.0008	.0009	.0010	.0012	.0013	.0015	.0017	.0019	.0021
18	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
19	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003	.0003	.0004
20	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002
21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

x	λ 8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0
0	.0003	.0003	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001
1	.0025	.0023	.0021	.0019	.0017	.0016	.0014	.0013	.0012	.0011
2	.0100	.0092	.0086	.0079	.0074	.0068	.0063	.0058	.0054	.0050
3	.0269	.0252	.0237	.0222	.0208	.0195	.0183	.0171	.0160	.0150
4	.0544	.0517	.0491	.0466	.0443	.0420	.0398	.0377	.0357	.0337
5	.0882	.0849	.0816	.0784	.0752	.0722	.0692	.0663	.0635	.0607
6	.1191	.1160	.1128	.1097	.1066	.1034	.1003	.0972	.0941	.0911
7	.1378	.1358	.1338	.1317	.1294	.1271	.1247	.1222	.1197	.1171
8	.1395	.1392	.1388	.1382	.1375	.1366	.1356	.1344	.1332	.1318
9	.1256	.1269	.1280	.1290	.1299	.1306	.1311	.1315	.1317	.1318
10	.1017	.1040	.1063	.1084	.1104	.1123	.1140	.1157	.1172	.1186
11	.0749	.0776	.0802	.0828	.0853	.0878	.0902	.0925	.0948	.0970
12	.0505	.0530	.0555	.0579	.0604	.0629	.0654	.0679	.0703	.0728
13	.0315	.0334	.0354	.0374	.0395	.0416	.0438	.0459	.0481	.0504
14	.0182	.0196	.0210	.0225	.0240	.0256	.0272	.0289	.0306	.0324
15	.0098	.0107	.0116	.0126	.0136	.0147	.0158	.0169	.0182	.0194
16	.0050	.0055	.0060	.0066	.0072	.0079	.0086	.0093	.0101	.0109
17	.0024	.0026	.0029	.0033	.0036	.0040	.0044	.0048	.0053	.0058
18	.0011	.0012	.0014	.0015	.0017	.0019	.0021	.0024	.0026	.0029
19	.0005	.0005	.0006	.0007	.0008	.0009	.0010	.0011	.0012	.0014
20	.0002	.0002	.0002	.0003	.0003	.0004	.0004	.0005	.0005	.0006
21	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0002	.0003
22	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001

x	λ 9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10
0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0000
1	.0010	.0009	.0009	.0008	.0007	.0007	.0006	.0005	.0005	.0005
2	.0046	.0043	.0040	.0037	.0034	.0031	.0029	.0027	.0025	.0023
3	.0140	.0131	.0123	.0115	.0107	.0100	.0093	.0087	.0081	.0076
4	.0319	.0302	.0285	.0269	.0254	.0240	.0226	.0213	.0201	.0189
5	.0581	.0555	.0530	.0506	.0483	.0460	.0439	.0418	.0398	.0378
6	.0881	.0851	.0822	.0793	.0764	.0736	.0709	.0682	.0656	.0631
7	.1145	.1118	.1091	.1064	.1037	.1010	.0982	.0955	.0928	.0901
8	.1302	.1286	.1269	.1251	.1232	.1212	.1191	.1170	.1148	.1126
9	.1317	.1315	.1311	.1306	.1300	.1293	.1284	.1274	.1263	.1251

تابع الجدول (٢ ـ ب)

x	λ 9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10
10	.1198	.1210	.1219	.1228	.1235	.1241	.1245	.1249	.1250	.1251
11	.0991	.1012	.1031	.1049	.1067	.1083	.1098	.1112	.1125	.1137
12	.0752	.0776	.0799	.0822	.0844	.0866	.0888	.0908	.0928	.0948
13	.0526	.0549	.0572	.0594	.0617	.0640	.0662	.0685	.0707	.0729
14	.0342	.0361	.0380	.0399	.0419	.0439	.0459	.0479	.0500	.0521
15	.0208	.0221	.0235	.0250	.0265	.0281	.0297	.0313	.0330	.0347
16	.0118	.0127	.0137	.0147	.0157	.0168	.0180	.0192	.0204	.0217
17	.0063	.0069	.0075	.0081	.0088	.0095	.0103	.0111	.0119	.0128
18	.0032	.0035	.0039	.0042	.0046	.0051	.0055	.0060	.0065	.0071
19	.0015	.0017	.0019	.0021	.0023	.0026	.0028	.0031	.0034	.0037
20	.0007	.0008	.0009	.0010	.0011	.0012	.0014	.0015	.0017	.0019
21	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
22	.0001	.0001	.0002	.0002	.0002	.0002	.0003	.0003	.0004	.0004
23	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002
24	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

x	λ 11	12	13	14	15	16	17	18	19	20
0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2	.0010	.0004	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
3	.0037	.0018	.0008	.0004	.0002	.0001	.0000	.0000	.0000	.0000
4	.0102	.0053	.0027	.0013	.0006	.0003	.0001	.0001	.0000	.0000
5	.0224	.0127	.0070	.0037	.0019	.0010	.0005	.0002	.0001	.0001
6	.0411	.0255	.0152	.0087	.0048	.0026	.0014	.0007	.0004	.0002
7	.0646	.0437	.0281	.0174	.0104	.0060	.0034	.0018	.0010	.0005
8	.0888	.0655	.0457	.0304	.0194	.0120	.0072	.0042	.0024	.0013
9	.1085	.0874	.0661	.0473	.0324	.0213	.0135	.0083	.0050	.0029
10	.1194	.1048	.0859	.0663	.0486	.0341	.0230	.0150	.0095	.0058
11	.1194	.1144	.1015	.0844	.0663	.0496	.0355	.0245	.0164	.0106
12	.1094	.1144	.1099	.0984	.0829	.0661	.0504	.0368	.0259	.0176
13	.0926	.1056	.1099	.1060	.0956	.0814	.0658	.0509	.0378	.0271
14	.0728	.0905	.1021	.1060	.1024	.0930	.0800	.0655	.0514	.0387
15	.0534	.0724	.0885	.0989	.1024	.0992	.0906	.0786	.0650	.0516
16	.0367	.0543	.0719	.0866	.0960	.0992	.0963	.0884	.0772	.0646
17	.0237	.0383	.0550	.0713	.0847	.0934	.0963	.0936	.0863	.0760
18	.0145	.0256	.0397	.0554	.0706	.0830	.0909	.0936	.0911	.0844
19	.0084	.0161	.0272	.0409	.0557	.0699	.0814	.0887	.0911	.0888
20	.0046	.0097	.0177	.0286	.0418	.0559	.0692	.0798	.0866	.0888
21	.0024	.0055	.0109	.0191	.0299	.0426	.0560	.0684	.0783	.0846
22	.0012	.0030	.0065	.0121	.0204	.0310	.0433	.0560	.0676	.0769
23	.0006	.0016	.0037	.0074	.0133	.0216	.0320	.0438	.0559	.0669
24	.0003	.0008	.0020	.0043	.0083	.0144	.0226	.0328	.0442	.0557
25	.0001	.0004	.0010	.0024	.0050	.0092	.0154	.0237	.0336	.0446
26	.0000	.0002	.0005	.0013	.0029	.0057	.0101	.0164	.0246	.0343
27	.0000	.0001	.0002	.0007	.0016	.0034	.0063	.0109	.0173	.0254
28	.0000	.0000	.0001	.0003	.0009	.0019	.0038	.0070	.0117	.0181
29	.0000	.0000	.0001	.0002	.0004	.0011	.0023	.0044	.0077	.0125
30	.0000	.0000	.0000	.0001	.0002	.0006	.0013	.0026	.0049	.0083
31	.0000	.0000	.0000	.0000	.0001	.0003	.0007	.0015	.0030	.0054
32	.0000	.0000	.0000	.0000	.0001	.0001	.0004	.0009	.0018	.0034
33	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0010	.0020
34	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0012
35	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0007
36	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004
37	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
38	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
39	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

الجدول (٤ ـ ب) : التوزيع الطبيعي

$$F(Z) = P_r(Z \le z) = \int_{-\infty}^{z} f(z)\,dz$$

$$Z \sim N(0,1), \; -\infty < Z < \infty$$

Z	0	1	2	3	4	5	6	7	8	9
−3.0	.0013	.0010	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000
−2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0238	.0233
−1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0300	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0570	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
− .9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
− .8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
− .7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
− .6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
− .5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
− .4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
− .3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
− .2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
− .1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
− 0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

تابع الجدول (٤ ـ ب)

Z	0	1	2	3	4	5	6	7	8	9
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9430	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9648	.9656	.9664	.9671	.9678	.9686	.9693	.9700	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9762	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9874	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.	.9987	.9990	.9993	.9995	.9997	.9998	.9998	.9999	.9999	1.0000

الجدول (٥ ـ ب) : توزيع بيتا

$$F(x) = P_r(X \le x) = \int_0^x f(u)\,du = 0{\cdot}05$$

$$X \sim \text{Beta}(\alpha, \beta)\,;\, 0 < x < 1$$

α \ β	·5	·0	·5	1·0	1·5	2·0	2·5	3·0	3·5
·5	·0061558	·0025000	·0015429	·0011119	·0³86320	·0³71179	·0³60300	·0³52300	·0³46170
·0	·097500	·050000	·033617	·025321	·020308	·016952	·014548	·012741	·011334
·5	·22852	·13572	·097308	·076010	·062413	·052962	·046007	·040671	·036447
1·0	·34163	·22361	·16825	·13535	·11339	·097611	·085727	·076440	·068979
1·5	·43074	·30171	·23553	·19403	·16528	·14409	·12778	·11482	·10427
2·0	·50053	·36840	·29599	·24860	·21477	·18926	·16927	·15316	·13989
2·5	·55593	·42489	·34929	·29811	·26063	·23182	·20890	·19019	·17461
3·0	·60071	·47237	·39607	·34259	·30260	·27134	·24613	·22532	·20783
3·5	·63751	·51390	·43716	·38246	·34080	·30777	·28082	·25835	·23930
4·0	·66824	·54928	·47338	·41820	·37553	·34126	·31301	·28924	·26894
4·5	·69425	·58003	·50546	·45032	·40712	·37203	·34283	·31807	·29677
5·0	·71654	·60696	·53402	·47930	·43590	·40031	·37044	·34494	·32286
5·5	·73683	·63073	·55958	·50551	·46219	·42635	·39604	·37000	·34732
6·0	·75268	·65184	·58256	·52932	·48626	·45036	·41980	·39338	·37025
6·5	·76754	·37070	·60333	·55102	·50836	·47255	·44187	·41521	·39176
7·0	·78072	·68736	·62217	·57036	·52872	·49310	·46242	·43563	·41196
7·5	·79249	·70297	·63933	·58907	·54750	·51217	·48159	·45474	·43094
8·0	·80307	·71687	·65503	·60584	·53490	·52991	·49949	·47267	·44880
8·5	·81263	·72954	·66944	·62131	·58103	·54645	·51624	·48951	·46564
9·0	·82131	·74113	·68271	·63564	·59605	·56189	·53194	·50535	·48152
9·5	·82923	·75178	·69496	·64894	·61004	·57655	·54669	·52027	·49652
10·0	·83647	·76160	·70632	·66132	·62312	·58990	·56056	·53434	·51071
10·5	·84313	·77067	·71637	·67287	·63536	·60253	·57363	·54764	·52415
11·0	·84927	·77908	·72669	·63366	·64684	·61461	·58596	·56022	·53689
11·5	·85494	·78690	·73586	·69377	·65764	·62590	·59761	·57213	·54898
12·0	·26021	·79418	·74444	·70327	·66780	·63659	·60864	·58343	·56048
12·5	·86511	·80099	·75249	·71219	·37738	·34663	·61909	·59416	·57141
13·0	·86967	·80736	·76004	·72060	·68643	·65617	·62900	·60436	·58183
13·5	·87394	·81334	·76715	·72854	·69499	·36522	·63842	·61407	·59177
14·0	·87794	·81896	·77386	·73604	·70311	·67381	·64738	·62332	·60125
19·0	·90724	·86089	·82447	·79327	·76559	·74053	·71758	·69636	·67663
20·0	·93748	·90497	·87881	·85591	·82617	·81606	·79824	·78150	·76569
50·0	·96837	·95130	·93720	·92458	·91290	·90192	·89148	·88150	·87191
∞	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000

ملاحظة ، ان كل تعني وان كل λ تعني

تابع الجدول (٥ ـ ب)

β^* / α^*	4·0	5·0	6·5	9·0	11·0	14·0	19·0	29·0	59·0
− ·5	$\cdot 0^{3}41325$	$\cdot 0^{3}34154$	$\cdot 0^{3}27098$	$\cdot 0^{3}20156$	$\cdot 0^{3}16727$	$\cdot 0^{3}13326$	$\cdot 0^{4}99535$	$\cdot 0^{4}66082$	$\cdot 0^{4}32904$
·0	·010206	·0085124	·0068158	·0051162	·0042653	·0034137	·0025614	·0017083	$\cdot 0^{3}85452$
·5	·033020	·027794	·022465	·017026	·014264	·011472	·0086511	·0057991	·0029157
1·0	·062860	·053376	·043541	·033319	·028053	·022679	·017191	·011585	·0058568
1·5	·095510	·081790	·067312	·051995	·043994	·035747	·027240	·018458	·0093841
2·0	·12876	·11111	·092207	·071870	·061103	·049898	·038224	·026043	·013317
2·5	·16142	·14029	·11733	·092238	·078783	·064651	·049781	·034103	·017540
3·0	·19290	·16875	·14216	·11267	·096658	·079695	·061676	·042481	·021976
3·5	·22292	·19618	·16638	·13288	·11449	·094827	·073748	·051068	·026572
4·0	·25137	·22244	·18984	·15272	·13211	·10991	·085885	·059786	·031288
4·5	·27823	·24746	·21244	·17207	·14943	·12484	·098003	·068575	·036094
5·0	·30354	·27125	·23413	·19086	·16636	·13955	·11006	·077394	·040967
5·5	·32737	·29383	·25492	·20908	·18288	·15401	·12199	·086209	·045889
6·0	·34981	·31524	·27481	·22669	·19895	·16818	·13377	·094994	·050847
6·5	·37095	·33554	·29382	·24370	·21457	·18203	·14539	·10373	·055827
7·0	·39086	·35480	·31199	·26011	·22972	·19556	·15682	·11240	·060821
7·5	·40965	·37307	·32936	·27594	·24441	·20877	·16805	·12090	·065820
8·0	·42738	·39041	·34596	·29120	·25865	·22164	·17908	·12950	·070818
8·5	·44414	·40689	·36183	·30591	·27244	·23418	·18989	·13791	·075809
9·0	·45999	·42256	·37701	·32009	·28580	·24639	·20050	·14622	·080789
9·5	·47501	·43746	·39154	·33375	·29874	·25828	·21088	·15442	·085753
10·0	·48925	·45165	·40544	·34693	·31126	·26985	·22106	·16252	·090698
10·5	·50276	·46518	·41877	·35964	·32340	·28112	·23102	·17051	·095622
11·0	·51560	·47808	·43154	·37190	·33515	·29208	·24077	·17838	·10052
11·5	·52782	·49040	·44379	·38373	·34653	·30275	·25032	·18615	·10539
12·0	·53945	·50217	·45554	·39516	·35756	·31314	·25966	·19379	·11024
12·5	·55054	·51343	·46683	·40619	·36826	·32325	·26880	·20133	·11505
13·0	·56112	·52420	·47768	·41685	·37862	·33309	·27775	·20875	·11983
13·5	·57122	·53452	·48812	·42715	·38867	·34267	·28650	·21606	·12458
14·0	·58088	·54442	·49816	·43711	·39842	·35200	·29507	·22326	·12930
19·0	·65819	·62459	·58083	·52099	·48175	·43321	·37136	·28936	·17453
29·0	·75070	·72282	·68535	·63185	·59522	·54807	·48477	·39458	·25416
59·0	·86266	·84504	·82047	·78342	·75661	·72016	·66738	·58326	·42519
∞	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000	1·00000

الجدول (٦ ـ ب) : توزيع مربع كاي

$$F(\chi^2) = P_r(\chi^2 \le \chi^2_n(\alpha)) = 1 - \alpha$$

$$\chi^2 \sim \chi^2_{(n)}, 0 < \chi^2 < \infty$$

n \ α	·005	·01	·025	·05	·10	·25	·50
1	392704.10^{-10}	157088.10^{-9}	982069.10^{-9}	393214.10^{-8}	·0157908	·1015308	·454936
2	·0100251	·0201007	·0506356	·102587	·210721	·575364	1·38629
3	·0717218	·114832	·215795	·351846	·584374	1·212534	2·36597
4	·206989	·297109	·484419	·710723	1·063623	1·92256	3·35669
5	·411742	·554298	·831212	1·145476	1·61031	2·67460	4·35146
6	·675727	·872090	1·23734	1·63538	2·20413	3·45460	5·34812
7	·989256	1·239043	1·68987	2·16735	2·83311	4·25485	6·34581
8	1·34441	1·64650	2·17973	2·73264	3·48954	5·07064	7·34412
9	1·73493	2·08790	2·70039	3·32511	4·16816	5·89883	8·34283
10	2·15586	2·55821	3·24697	3·94030	4·86518	6·73720	9·34182
11	2·60322	3·05348	3·81575	4·57481	5·57778	7·58414	10·3410
12	3·07382	3·57057	4·40379	5·22603	6·30380	8·43842	11·3403
13	3·56503	4·10692	5·00875	5·89186	7·04150	9·29907	12·3398
14	4·07467	4·66043	5·62873	6·57063	7·78953	10·1653	13·3393
15	4·60092	5·22935	6·26214	7·26094	8·54676	11·0365	14·3389
16	5·14221	5·81221	6·90766	7·96165	9·31224	11·9122	15·3385
17	5·69722	6·40776	7·56419	8·67176	10·0852	12·7919	16·3382
18	6·26480	7·01491	8·23075	9·39046	10·8649	13·6753	17·3379
19	6·84397	7·63273	8·90652	10·1170	11·6509	14·5620	18·3377
20	7·43384	8·26040	9·59078	10·8508	12·4426	15·4518	19·3374
21	8·03365	8·89720	10·28293	11·5913	13·2396	16·3444	20·3372
22	8·64272	9·54249	10·9823	12·3380	14·0415	17·2396	21·3370
23	9·26043	10·19567	11·6886	13·0905	14·8480	18·1373	22·3369
24	9·88623	10·8564	12·4012	13·8484	15·6587	19·0373	23·3367
25	10·5197	11·5240	13·1197	14·6114	16·4734	19·9393	24·3366
26	11·1602	12·1981	13·8439	15·3792	17·2919	20·8434	25·3365
27	11·8076	12·8785	14·5734	16·1514	18·1139	21·7494	26·3363
28	12·4613	13·5647	15·3079	16·9279	18·9392	22·6572	27·3362
29	13·1211	14·2565	16·0471	17·7084	19·7677	23·5666	28·3361
30	13·7867	14·9535	16·7908	18·4927	20·5992	24·4776	29·3360
40	20·7065	22·1643	24·4330	26·5093	29·0505	33·6603	39·3353
50	27·9907	29·7067	32·3574	34·7643	37·6886	42·9421	49·3349
60	35·5345	37·4849	40·4817	43·1880	46·4589	52·2938	59·3347
70	43·2752	45·4417	48·7576	51·7393	55·3289	61·6983	69·3345
80	51·1719	53·5401	57·1532	60·3915	64·2778	71·1445	79·3343
90	59·1963	61·7541	65·6466	69·1260	73·2911	80·6247	89·3342
100	67·3276	70·0649	74·2219	77·9295	82·3581	90·1332	99·3341

تابع الجدول (٦ ــ ب) :

n \ α	·75	·90	·95	·975	·99	·995	·999
1	1·32330	2·70554	3·84146	5·02389	6·63490	7·87944	10·828
2	2·77259	4·60517	5·99146	7·37776	9·21034	10·5966	13·816
3	4·10834	6·25139	7·81473	9·34840	11·3449	12·8382	16·266
4	5·38527	7·77944	9·48773	11·1433	13·2767	14·8603	18·467
5	6·62568	9·23636	11·0705	12·8325	15·0863	16·7496	20·515
6	7·84080	10·6446	12·5916	14·4494	16·8119	18·5476	22·458
7	9·03715	12·0170	14·0671	16·0128	18·4753	20·2777	24·322
8	10·2189	13·3616	15·5073	17·5345	20·0902	21·9550	26·125
9	11·3888	14·6837	16·9190	19·0228	21·6660	23·5894	27·877
10	12·5489	15·9872	18·3070	20·4832	23·2093	25·1882	29·588
11	13·7007	17·2750	19·6751	21·9200	24·7250	26·7568	31·264
12	14·8454	18·5493	21·0261	23·3367	26·2170	28·2995	32·909
13	15·9839	19·8119	22·3620	24·7356	27·6882	29·8195	34·528
14	17·1169	21·0641	23·6848	26·1189	29·1412	31·3194	36·123
15	18·2451	22·3071	24·9958	27·4884	30·5779	32·8013	37·697
16	19·3689	23·5418	26·2962	28·8454	31·9999	34·2672	39·252
17	20·4887	24·7690	27·5871	30·1910	33·4087	35·7185	40·790
18	21·6049	25·9894	28·8693	31·5264	34·8053	37·1565	42·312
19	22·7178	27·2036	30·1435	32·8523	36·1909	38·5823	43·820
20	23·8277	28·4120	31·4104	34·1696	37·5662	39·9968	45·315
21	24·9348	29·6151	32·6706	35·4789	38·9322	41·4011	46·797
22	26·0393	30·8133	33·9244	36·7807	40·2894	42·7957	48·268
23	27·1413	32·0069	25·1725	38·0756	41·6384	44·1813	49·728
24	28·2412	33·1962	36·4150	39·3641	42·9798	45·5585	51·179
25	29·3389	34·3816	37·6525	40·6465	44·3141	46·9279	52·618
26	30·4346	35·5632	38·8851	41·9232	45·6417	48·2899	54·052
27	31·5284	36·7412	40·1133	43·1945	46·9629	49·6449	55·476
28	32·6205	37·9159	41·3371	44·4608	48·2782	50·9934	56·892
29	33·7109	39·0875	42·5570	45·7223	49·5879	52·3356	58·301
30	34·7997	40·2560	43·7730	46·9792	50·8922	53·6720	59·703
40	45·6160	51·8051	55·7585	59·3417	63·6907	66·7660	73·402
50	56·3336	63·1671	67·5048	71·4202	76·1539	79·4900	86·661
60	66·9815	74·3970	79·0819	83·2977	88·3794	91·9517	99·607
70	77·5767	85·5270	90·5312	95·0232	100·425	104·215	112·317
80	88·1303	96·5782	101·879	106·629	112·329	116·321	124·839
90	98·6499	107·565	113·145	118·136	124·116	128·299	137·208
100	109·141	118·498	124·342	129·561	135·807	140·169	149·449

الجدول : (٧ ـ ب) توزيع t

$$F(t) = P_r(t \leq t_n(\alpha) = 1 - \alpha$$

$$t \sim t_{(n)} ; -\infty < t < \infty$$

n \ α	·60	·75	·90	·95	·975	·99	·995	·9975	·999	·9995
1	·325	1·000	3·078	6·314	12·706	31·821	63·657	127·32	318·31	636·62
2	·289	·816	1·886	2·920	4·303	6·965	9·925	14·089	22·327	31·598
3	·277	·765	1·638	2·353	3·182	4·541	5·841	7·453	10·214	12·924
4	·271	·741	1·533	2·132	2·776	3·747	4·604	5·598	7·173	8·610
5	·267	·727	1·476	2·015	2·571	3·365	4·032	4·773	5·893	6·869
6	·265	·718	1·440	1·943	2·447	3·143	3·707	4·317	5·208	5·959
7	·263	·711	1·415	1·895	2·365	2·998	3·499	4·029	4·785	5·408
8	·262	·706	1·397	1·860	2·306	2·896	3·355	3·833	4·501	5·041
9	·261	·703	1·383	1·833	2·262	2·821	3·250	3·690	4·297	4·781
10	·260	·700	1·372	1·812	2·228	2·764	3·169	3·581	4·144	4·587
11	·260	·697	1·363	1·796	2·201	2·718	3·106	3·497	4·025	4·437
12	·259	·695	1·356	1·782	2·179	2·681	3·055	3·428	3·930	4·318
13	·259	·694	1·350	1·771	2·160	2·650	3·012	3·372	3·852	4·221
14	·258	·692	1·345	1·761	2·145	2·624	2·977	3·326	3·787	4·140
15	·258	·691	1·341	1·753	2·131	2·602	2·947	3·286	3·733	4·073
16	·258	·690	1·337	1·746	2·120	2·583	2·921	3·252	3·686	4·015
17	·257	·689	1·333	1·740	2·110	2·567	2·898	3·222	3·646	3·965
18	·257	·688	1·330	1·734	2·101	2·552	2·878	3·197	3·610	3·922
19	·257	·688	1·328	1·729	2·093	2·539	2·861	3·174	3·579	3·883
20	·257	·687	1·325	1·725	2·086	2·528	2·845	3·153	3·552	3·850
21	·257	·686	1·323	1·721	2·080	2·518	2·831	3·135	3·527	3·819
22	·256	·686	1·321	1·717	2·074	2·508	2·819	3·119	3·505	3·792
23	·256	·685	1·319	1·714	2·069	2·500	2·807	3·104	3·485	3·767
24	·256	·685	1·318	1·711	2·064	2·492	2·797	3·091	3·467	3·745
25	·256	·684	1·316	1·708	2·060	2·485	2·787	3·078	3·450	3·725
26	·256	·684	1·315	1·706	2·056	2·479	2·779	3·067	3·435	3·707
27	·256	·684	1·314	1·703	2·052	2·473	2·771	3·057	3·421	3·690
28	·256	·683	1·313	1·701	2·048	2·467	2·763	3·047	3·408	3·674
29	·256	·683	1·311	1·699	2·045	2·462	2·756	3·038	3·396	3·659
30	·256	·683	1·310	1·697	2·042	2·457	2·750	3·030	3·385	3·646
40	·255	·681	1·303	1·684	2·021	2·423	2·704	2·971	3·307	3·551
60	·254	·679	1·296	1·671	2·000	2·390	2·660	2·915	3·232	3·460
120	·254	·677	1·289	1·658	1·980	2·358	2·617	2·860	3·160	3·373
∞	·253	·674	1·282	1·645	1·960	2·326	2·576	2·807	3·090	3·291

الجدول (٨ ــ ب) : توزيع F

$$G(f) = P_r(f \le f_{n_1, n_2}(0{\cdot}05)) = 0{\cdot}95$$

$$f \sim F(n_1, n_2) ; 0 < f < \infty$$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	240·5	241·9	243·9	245·9	248·0	249·1	250·1	251·1	252·2	253·3	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·38	19·40	19·41	19·43	19·45	19·45	19·46	19·47	19·48	19·49	19·50
3	10·13	9·55	9·28	9·12	9·01	8·94	8·89	8·85	8·81	8·79	8·74	8·70	8·66	8·64	8·62	8·59	8·57	8·55	8·53
4	7·71	6·94	6·59	6·39	6·26	6·16	6·09	6·04	6·00	5·96	5·91	5·86	5·80	5·77	5·75	5·72	5·69	5·66	5·63
5	6·61	5·79	5·41	5·19	5·05	4·95	4·88	4·82	4·77	4·74	4·68	4·62	4·56	4·53	4·50	4·46	4·43	4·40	4·36
6	5·99	5·14	4·76	4·53	4·39	4·28	4·21	4·15	4·10	4·06	4·00	3·94	3·87	3·84	3·81	3·77	3·74	3·70	3·67
7	5·59	4·74	4·35	4·12	3·97	3·87	3·79	3·73	3·68	3·64	3·57	3·51	3·44	3·41	3·38	3·34	3·30	3·27	3·23
8	5·32	4·46	4·07	3·84	3·69	3·58	3·50	3·44	3·39	3·35	3·28	3·22	3·15	3·12	3·08	3·04	3·01	2·97	2·93
9	5·12	4·26	3·86	3·63	3·48	3·37	3·29	3·23	3·18	3·14	3·07	3·01	2·94	2·90	2·86	2·83	2·79	2·75	2·71
10	4·96	4·10	3·71	3·48	3·33	3·22	3·14	3·07	3·02	2·98	2·91	2·85	2·77	2·74	2·70	2·66	2·62	2·58	2·54
11	4·84	3·98	3·59	3·36	3·20	3·09	3·01	2·95	2·90	2·85	2·79	2·72	2·65	2·61	2·57	2·53	2·49	2·45	2·40
12	4·75	3·89	3·49	3·26	3·11	3·00	2·91	2·85	2·80	2·75	2·69	2·62	2·54	2·51	2·47	2·43	2·38	2·34	2·30
13	4·67	3·81	3·41	3·18	3·03	2·92	2·83	2·77	2·71	2·67	2·60	2·53	2·46	2·42	2·38	2·34	2·30	2·25	2·21
14	4·60	3·74	3·34	3·11	2·96	2·85	2·76	2·70	2·65	2·60	2·53	2·46	2·39	2·35	2·31	2·27	2·22	2·18	2·13
15	4·54	3·68	3·29	3·06	2·90	2·79	2·71	2·64	2·59	2·54	2·48	2·40	2·33	2·29	2·25	2·20	2·16	2·11	2·07
16	4·49	3·63	3·24	3·01	2·85	2·74	2·66	2·59	2·54	2·49	2·42	2·35	2·28	2·24	2·19	2·15	2·11	2·06	2·01
17	4·45	3·59	3·20	2·96	2·81	2·70	2·61	2·55	2·49	2·45	2·38	2·31	2·23	2·19	2·15	2·10	2·06	2·01	1·96
18	4·41	3·55	3·16	2·93	2·77	2·66	2·58	2·51	2·46	2·41	2·34	2·27	2·19	2·15	2·11	2·06	2·02	1·97	1·92
19	4·38	3·52	3·13	2·90	2·74	2·63	2·54	2·48	2·42	2·38	2·31	2·23	2·16	2·11	2·07	2·03	1·98	1·93	1·88
20	4·35	3·49	3·10	2·87	2·71	2·60	2·51	2·45	2·39	2·35	2·28	2·20	2·12	2·08	2·04	1·99	1·95	1·90	1·84
21	4·32	3·47	3·07	2·84	2·68	2·57	2·49	2·42	2·37	2·32	2·25	2·18	2·10	2·05	2·01	1·96	1·92	1·87	1·81
22	4·30	3·44	3·05	2·82	2·66	2·55	2·46	2·40	2·34	2·30	2·23	2·15	2·07	2·03	1·98	1·94	1·89	1·84	1·78
23	4·28	3·42	3·03	2·80	2·64	2·53	2·44	2·37	2·32	2·27	2·20	2·13	2·05	2·01	1·96	1·91	1·86	1·81	1·76
24	4·26	3·40	3·01	2·78	2·62	2·51	2·42	2·36	2·30	2·25	2·18	2·11	2·03	1·98	1·94	1·89	1·84	1·79	1·73
25	4·24	3·39	2·99	2·76	2·60	2·49	2·40	2·34	2·28	2·24	2·16	2·09	2·01	1·96	1·92	1·87	1·82	1·77	1·71
26	4·23	3·37	2·98	2·74	2·59	2·47	2·39	2·32	2·27	2·22	2·15	2·07	1·99	1·95	1·90	1·85	1·80	1·75	1·69
27	4·21	3·35	2·96	2·73	2·57	2·46	2·37	2·31	2·25	2·20	2·13	2·06	1·97	1·93	1·88	1·84	1·79	1·73	1·67
28	4·20	3·34	2·95	2·71	2·56	2·45	2·36	2·29	2·24	2·19	2·12	2·04	1·96	1·91	1·87	1·82	1·77	1·71	1·65
29	4·18	3·33	2·93	2·70	2·55	2·43	2·35	2·28	2·22	2·18	2·10	2·03	1·94	1·90	1·85	1·81	1·75	1·70	1·64
30	4·17	3·32	2·92	2·69	2·53	2·42	2·33	2·27	2·21	2·16	2·09	2·01	1·93	1·89	1·84	1·79	1·74	1·68	1·62
40	4·08	3·23	2·84	2·61	2·45	2·34	2·25	2·18	2·12	2·08	2·00	1·92	1·84	1·79	1·74	1·69	1·64	1·58	1·51
60	4·00	3·15	2·76	2·53	2·37	2·25	2·17	2·10	2·04	1·99	1·92	1·84	1·75	1·70	1·65	1·59	1·53	1·47	1·39
120	3·92	3·07	2·68	2·45	2·29	2·17	2·09	2·02	1·96	1·91	1·83	1·75	1·66	1·61	1·55	1·50	1·43	1·35	1·25
∞	3·84	3·00	2·60	2·37	2·21	2·10	2·01	1·94	1·88	1·83	1·75	1·67	1·57	1·52	1·46	1·39	1·32	1·22	1·00

تابع الجدول (٨ ــ ب)

$$G(f) = P_r(f \le f_{n_1, n_2}(0.01)) = 0.99$$

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	5000	5403	5625	5764	5859	5928	5982	6023	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	3.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

الجدول (٩ ـ ب) : توزيع المدى القياسي

$P_r(S \leq S_{n,m}(0.05)) = 0.95$

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96	53.20	54.33	55.36	56.32	57.22	58.04	58.83	59.56
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.37	16.57	16.77
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95	10.15	10.35	10.52	10.69	10.84	10.98	11.11	11.24
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.66	8.79	8.91	9.03	9.13	9.23
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.10	7.17
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83	5.93	6.03	6.11	6.19	6.27	6.34	6.40	6.47
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.98	6.06	6.13	6.20	6.27	6.33
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79	5.86	5.93	5.99	6.05	6.11
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.55	5.64	5.71	5.79	5.85	5.91	5.97	6.03
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.57	5.65	5.72	5.78	5.85	5.90	5.96
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90
17	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.54	5.61	5.67	5.73	5.79	5.84
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17	5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23	5.31	5.39	5.46	5.53	5.59	5.65	5.70	5.75
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32	5.38	5.44	5.49	5.55	5.59
30	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92	5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.47
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82	4.90	4.98	5.04	5.11	5.16	5.22	5.27	5.31	5.36
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24
120	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13
∞	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.68	4.74	4.80	4.85	4.89	4.93	4.97	5.01

تابع الجدول (٩ ـ ب)

$$P_r\left(S \leq S_{n,m}(0.01)\right) = 0.99$$

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90.03	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6	253.2	260.0	266.2	271.8	277.0	281.8	286.3	290.4	294.3	298.0
2	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69	32.59	33.40	34.13	34.81	35.43	36.00	36.53	37.03	37.50	37.95
3	8.26	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69	17.13	17.53	17.89	18.22	18.52	18.81	19.07	19.32	19.55	19.77
4	6.51	8.12	9.17	9.96	10.58	11.10	11.55	11.93	12.27	12.57	12.84	13.09	13.32	13.53	13.73	13.91	14.08	14.24	14.40
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.48	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.36	8.55	8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65
8	4.75	5.04	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03	8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65	7.78	7.91	8.03	8.13	8.23	8.33	8.41	8.49	8.57
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.49	7.60	7.71	7.81	7.91	7.99	8.08	8.15	8.23
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19	7.27	7.35	7.42	7.48	7.55
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77	6.87	6.96	7.05	7.13	7.20	7.27	7.33	7.39
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73	6.81	6.87	6.94	7.00	7.05
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.65	6.73	6.79	6.85	6.91	6.97
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.28	6.37	6.45	6.52	6.59	6.65	6.71	6.77	6.82
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41
40	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69	5.76	5.83	5.90	5.96	6.02	6.07	6.12	6.16	6.21
60	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.78	5.84	5.89	5.93	5.97	6.01
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37	5.44	5.50	5.56	5.61	5.66	5.71	5.75	5.79	5.83
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65

الملحق (جـ)
مصطلحات رياضية واحصائية

- A -

Additive property	خاصية الجمع
Alternative hypothesis	فرضية بديلة
Analysis of variance	تحليل التباين
Applied mathematics	رياضيات تطبيقية
Approximation	تقريب
Arc - sine distribution	توزيع الجيب القوسي
Associative Law	قانون الترتيب
Asymptotic distribution	توزيع محاذي
Axioms	بديهيات

- B -

Bay's theorem	نظرية بيز
Bernoulli trials	محاولات برنولي
Best estimator	افضل تقدير
Beta distribution	توزيع بيتا
Binomial distribution	توزيع ثنائي الحدين
Binomial theorem	نظرية ثنائي الحدين

- C -

Cauchy distribution	توزيع كوشي
Central absolute moments	عزوم مطلقة مركزية
Central limit theorem	مبرهنة الغاية المركزية
Central moments	عزوم مركزية
Characteristic function	دالة مميزة (وصفية)
Chebyshev's inequality	متباينة تشيبيشيف
Chi - square distribution	توزيع مربع كاي
Coefficient of dispersion	معامل التشتت
Coefficient of skewness	معامل الالتواء

Coefficient of variation	معامل الاختلاف
Commutative law	قانون الابدال
Complement	متممة
Complex number	عدد معقد
Composite hypothesis	فرضية مركبة
Compound distributions	توزيعات مركبة
Concave function	دالة مقعره
Conditional	شرْطي
Conditional distribution	توزيع شرْطي
Conditional probability	احتمال شرْطي
Confidence interval	فترة ثقة
Confluent hypergeometric function	الدالة الزائدية المندمجة
Conjugate function	دالة مرافقة
Consistent estimator	تقدير متسق
Continuity correcttion	مصحح الاستمرارية
Continuous	مستمر
Convergency	تقارب
Convex function	دالة محدبة
Convolution formula	صيغة الالتفافية
Correlation coefficient	معامل ارتباط
Countable set	مجموعة قابله للعد
Covariance	تباين مشترك
Critical region	منطقة حرجة
Critical Values	قيم حرجة
Cumulant generating function	دالة مولدة تراكمية
Cumulative distribution function	دالة التوزيع التراكمية

- D -

Deciles	العُشيرات
Degrees of freedoms	درجات حرية
Density function	دالة كثافة

Digamma function	دالة كاما المضاعفة
Discontinuity	انقطاع (عدم الاستمرارية)
Discrete	متقطع (منفصل)
Divergency	تباعد
Domain	منطلق

- E -

Efficiency	كفاءه
Efficient estimator	تقدير كفوء
Element	عنصر
Empty set	مجموعة خالية
Equally likely events	حوادث ذات فرص متساوية
Equivalent sets	مجموعات متكافئة
Euler's constant	ثابت اويلر
Event	حادثة
Expected frequency	تكرار متوقع
Exponential distribution	التوزيع الاسي
Extreme value distribution	توزيع القيمة المتطرفة

- F -

Factorial moments	عزوم عاملية
Finite set	مجموعة منتهية
Flatness	تسطح
Fourier's inversion theorem	نظرية الانعكاس لـ فوراير

- G -

Gamma distribution	توزيع كاما
Geometric distribution	التوزيع الهندسي
Goodness of fit	حسن المطابقة
Gumbel distribution	توزيع كامبل

- H -

Harmonic mean	وسط توافقي
Hypergeometric distribution	توزيع هندسي زائدي

- I -

Idempotent law	قانون اللانمو
Identity matrix	مصفوفة احادية
Incomplete	غير تام (ناقص)
Inequality	متباينة (متراجحة)
Infinite set	مجموعة غير منتهية
Inflexion points	نقاط انقلاب
Intersection	تقاطع
Interval estimation	التقدير بفترة

- J -

Joint distribution	توزيع مشترك

- K -

Kurtosis	تفلطح

- L -

Laplace distribution	توزيع لابلاس
Law of Large numbers	قانون الاعداد الكبيرة
Level of significance	مستوى المعنوية
Likelihood function	دالة امكان
Limit theorems	نظريات الغاية
Limiting distribution	توزيع مقيد
Linear combination	تركيب خطي
Logistic distribution	التوزيع السوقي
Log normal distribution	التوزيع اللوغارتمي الطبيعي
Lôpital's rule	قاعدة لوبيتل

M

Maclaurin's expansion	مفكوك مكلورين
Maps	يُطبق

Marginal distribution	توزيع حدي (هامشي)
Mass function	دالة كتلة
Mathematical expectation	توقع رياضي
Mathematical model	نموذج رياضي
Maximum likelihood	امكان اعظم
Mean	متوسط (وسط)
Mean deviation	انحراف مطلق (متوسط)
Median	وسيط
Mid - range	منتصف المدى
Mixture of distributions	خلط التوزيعات
Mode	منوال
Moments	عزوم
Moments about the origin	عزوم حول نقطة الاصل
Moment generating function	دالة مولدة للعزوم
Most powerful test (M. P. T)	الاختبار الاكثر قوة
Multinomial distribution	توزيع متعدد الحدود
Multiple correlation	ارتباط متعدد
Multivariate distribution	توزيع متعدد المتغيرات
Mutually exclusive event	حوادث متنافية

- N -

Negative binomial distribution	توزيع ثنائي الحدين السالب
Non - central moments	عزوم لامركزية
Non - decreasing function	دالة غير متناقصة
Non - negative function	دالة غير سالبة
Normal distribution	التوزيع الطبيعي
Null hypothesis	فرضية العدم

- O -

Optimum test	اختبار امثل
Order	مرتبة
Order statistics	احصاءات مرتبة

Orthogonal	متعامد
Other wise	في غير ذلك (في احوال اخرى)

- P -

Parameter	معلمه
Parametric distribution	توزيع معلمي
Pareto distribution	توزيع باريتو
Partial correlation	ارتباط جزئي
Peakedness	تدبب
Pearsonian system	منظومة بيرسون
Point estimation	التقدير بنقطة
Poisson distribution	توزيع بواسون
Polar coordinates	احداثيات قطبية
Polya's distribution	توزيع بوليا
Population	مجتمع
Power of a test	قوة اختبار
Power series distribution	توزيع متسلسلة القوى
Probability curve	منحنى احتمالي
Probability distribution	توزيع احتمالي
Probability generating function	دالة مولدة احتمالية
Probability theory	نظرية الاحتمالات

- Q -

Quality control	الرقابة على الجودة
Quantiles	تجزئات
Quartiles	رُبيعات
Quartile deviation	انحراف ربيعي

- R -

Random sample	عينة عشوائية
Random variable	متغير عشوائي
Range	مدى

Rank	رتبة
Real - valued function	دالة ذات قيمة حقيقية
Recurrence formula	صيغة التراجع
Reliability	معولية

- S -

Sample point	نقطة عينة
Sample space	فضاء العينة
Sampling distribution	توزيع معاينة
Sampling techniques	اساليب معاينة
Sequential analysis	تحليل متسلسل
Set difference	فضلة المجموعة
Set theory	نظرية المجموعات
Single - valued function	دالة وحيدة القيمة
Simple hypothesis	فرضية بسيطة
Skewed distribution	توزيع ملتو
Skewness	الالتواء
Space	فضاء
Standard deviation	انحراف معياري
Statistic	مؤشر احصائي
Stochastic convergence	التقارب التصادفي
Stochastic independance	الاستقلال التصادفي
Studentized range	المدى القياسي
Sub set	مجموعة جزئية
Sufficient Statistic	مؤشر احصائي كافي
Symmetric	متماثل
Symmetric distribution	توزيع متماثل

- T -

Theory of estimation	نظرية التقدير
Testing hypotheses	اختبار الفرضيات
Transformation	تحويل
Truncated distribution	توزيع مقطوع (مبتور)

- U -

Unbiased estimator	تقدير غير متحيز
Uncorrelated	غير مرتبط
Uncountable set	مجموعة غير قابلة للعد
Uniform distribution	توزيع منتظم
Uniformly M. P. T.	الاختبار الاكثر قوة بانتظام
Union	اتحاد
Unique function	دالة وحيدة (فريدة)
Universal set	المجموعة الشاملة

- V -

Variance	تباين
Venn diagrams	مخططات ڤين

- W -

Wald distribution	توزيع والد
Weibull distribution	توزيع وايبل

رقم الايداع في المكتبة الوطنية ببغداد ٤٥١ لسنة ١٩٩٠

دار ابن الاثير للطباعة والنشر
جامعة الموصل